/9603/1302

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, C07K 14/27, A01H 1/00, 3/00	A3	(11) International Publication Number: WO 98/54214 (43) International Publication Date: 3 December 1998 (03.12.98)

(21) International Application Number: PCT/US98/10874

(22) International Filing Date: 28 May 1998 (28.05.98)

(30) Priority Data:
60/048,109   30 May 1997 (30.05.97)   US

(71) Applicants: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US). EDEN BIOSCIENCE CORPORATION [US/US]; 11816 North Creek Parkway N., Bothell, WA 98011-8205 (US).

(72) Inventors: LABY, Ronald, J.; 1411 Haddon Street, Houston, TX 77006 (US). WEI, Zhong-Min; 8230 125th Court, Kirkland, WA 98034 (US). BEER, Steven, V.; 211 Hudson Street, Ithaca, NY 14850 (US).

(74) Agents: GOLDMAN, Michael, L. et al.; Nixon, Hargrave, Devans & Doyle LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report: 4 March 1999 (04.03.99)

(54) Title: HYPERSENSITIVE RESPONSE ELICITOR FRAGMENTS AND USES THEREOF

(57) Abstract

The present invention is directed to isolated fragments of an *Erwinia* hypersensitive response elicitor protein or polypeptide which fragments elicit a hypersensitive response in plants. Also disclosed are isolated DNA molecules which encode the *Erwinia* hypersensitive response eliciting fragment. Isolated fragments of hypersensitive response elicitor proteins or polypeptides, which elicit a hypersensitive response, and the isolated DNA molecules that encode them can be used to impart disease resistance to plants, to enhance plant growth, and/or to control insects on plants. This can be achieved by applying the hypersensitive response eliciting fragments in a non-infectious form to plants or plant seeds under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds. Alternatively, transgenic plants or plant seeds transformed with a DNA molecule encoding a hypersensitive response eliciting fragment can be provided and the transgenic plants or plants resulting from the transgenic plant seeds are grown under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds.

HYPERSENSITIVE RESPONSE ELICITOR FRAGMENTS AND USES THEREOF

This application claims benefit of U.S. Patent Application Serial

5   No. 60/048,109.

FIELD OF THE INVENTION

The present invention relates to fragments of a hypersensitive response elicitor

10   which fragments elicit a hypersensitive response and uses thereof.

BACKGROUND OF THE INVENTION

Interactions between bacterial pathogens and their plant hosts generally fall

15   into two categories: (1) compatible (pathogen-host), leading to intercellular bacterial

growth, symptom development, and disease development in the host plant; and (2)

incompatible (pathogen-nonhost), resulting in the hypersensitive response, a particular

type of incompatible interaction occurring, without progressive disease symptoms.

During compatible interactions on host plants, bacterial populations increase

20   dramatically and progressive symptoms occur. During incompatible interactions,

bacterial populations do not increase, and progressive symptoms do not occur.

The hypersensitive response is a rapid, localized necrosis that is associated

with the active defense of plants against many pathogens (Kiraly, Z., "Defenses

Triggered by the Invader: Hypersensitivity," pages 201-224 in: Plant Disease: An

25   Advanced Treatise, Vol. 5, J.G. Horsfall and E.B. Cowling, ed. Academic Press New

York (1980); Klement, Z., "Hypersensitivity," pages 149-177 in: Phytopathogenic

Prokaryotes, Vol. 2, M.S. Mount and G.H. Lacy, ed. Academic Press, New York

(1982)). The hypersensitive response elicited by bacteria is readily observed as a

tissue collapse if high concentrations ($\geq 10^7$ cells/ml) of a limited host-range pathogen

30   like Pseudomonas syringae or Erwinia amylovora are infiltrated into the leaves of

nonhost plants (necrosis occurs only in isolated plant cells at lower levels of

inoculum) (Klement, Z., "Rapid Detection of Pathogenicity of Phytopathogenic

Pseudomonads," Nature 199:299-300; Klement, et al., "Hypersensitive Reaction

Induced by Phytopathogenic Bacteria in the Tobacco Leaf," Phytopathology 54:474-

- 2 -

477 (1963); Turner, et al., "The Quantitative Relation Between Plant and Bacterial
Cells Involved in the Hypersensitive Reaction," Phytopathology 64:885-890 (1974);
Klement, Z., "Hypersensitivity," pages 149-177 in Phytopathogenic Prokaryotes,
Vol. 2., M.S. Mount and G.H. Lacy, ed. Academic Press, New York (1982)). The

5   capacities to elicit the hypersensitive response in a nonhost and be pathogenic in a
host appear linked. As noted by Klement, Z., "Hypersensitivity," pages 149-177 in
Phytopathogenic Prokaryotes, Vol. 2., M.S. Mount and G.H. Lacy, ed. Academic
Press, New York, these pathogens also cause physiologically similar, albeit delayed,
necroses in their interactions with compatible hosts. Furthermore, the ability to

10   produce the hypersensitive response or pathogenesis is dependent on a common set of
genes, denoted hrp (Lindgren, P.B., et al., "Gene Cluster of Pseudomonas syringae
pv. 'phaseolicola' Controls Pathogenicity of Bean Plants and Hypersensitivity on
Nonhost Plants," J. Bacteriol. 168:512-22 (1986); Willis, D.K., et al., "hrp Genes of
Phytopathogenic Bacteria," Mol. Plant-Microbe Interact. 4:132-138 (1991)).

15   Consequently, the hypersensitive response may hold clues to both the nature of plant
defense and the basis for bacterial pathogenicity.

The hrp genes are widespread in gram-negative plant pathogens, where they
are clustered, conserved, and in some cases interchangeable (Willis, D.K., et al., "hrp
Genes of Phytopathogenic Bacteria," Mol. Plant-Microbe Interact. 4:132-138 (1991);

20   Bonas, U., "hrp Genes of Phytopathogenic Bacteria," pages 79-98 in: Current Topics
in Microbiology and Immunology: Bacterial Pathogenesis of Plants and Animals -
Molecular and Cellular Mechanisms, J.L. Dangl, ed. Springer-Verlag, Berlin (1994)).
Several hrp genes encode components of a protein secretion pathway similar to one
used by Yersinia, Shigella, and Salmonella spp. to secrete proteins essential in animal

25   diseases (Van Gijsegem, et al., "Evolutionary Conservation of Pathogenicity
Determinants Among Plant and Animal Pathogenic Bacteria," Trends Microbiol.
1:175-180 (1993)). In E. amylovora, P. syringae, and P. solanacearum, hrp genes
have been shown to control the production and secretion of glycine-rich, protein
elicitors of the hypersensitive response (He, S.Y., et al. "Pseudomonas Syringae pv.

30   Syringae HarpinPss: a Protein that is Secreted via the Hrp Pathway and Elicits the
Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), Wei, Z.-H., et al.,
"HrpI of Erwinia amylovora Functions in Secretion of Harpin and is a Member of a

- 3 -

New Protein Family," J. Bacteriol. 175:7958-7967 (1993); Arlat, M. et al. "PopA1, a Protein Which Induces a Hypersensitive-like Response on Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum,*" EMBO J. 13:543-553 (1994)).

5   The first of these proteins was discovered in *E. amylovora* Ea321, a bacterium that causes fire blight of rosaceous plants, and was designated harpin (Wei, Z.-M., et al, "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora,*" Science 257:85-88 (1992)). Mutations in the encoding *hrpN* gene revealed that harpin is required for *E. amylovora* to elicit a hypersensitive

10   response in nonhost tobacco leaves and incite disease symptoms in highly susceptible pear fruit. The *P. solanacearum* GMI1000 PopA1 protein has similar physical properties and also elicits the hypersensitive response in leaves of tobacco, which is not a host of that strain (Arlat, et al. "PopA1, a Protein Which Induces a Hypersensitive-like Response on Specific Petunia Genotypes, is Secreted via the Hrp

15   Pathway of *Pseudomonas solanacearum,*" EMBO J. 13:543-53 (1994)). However, *P. solanacearum popA* mutants still elicit the hypersensitive response in tobacco and incite disease in tomato. Thus, the role of these glycine-rich hypersensitive response elicitors can vary widely among gram-negative plant pathogens.

Other plant pathogenic hypersensitive response elicitors have been isolated,

20   cloned, and sequenced. These include: *Erwinia chrysanthemi* (Bauer, et. al., "*Erwinia chrysanthemi* Harpin$_{Ech}$: Soft-Rot Pathogenesis," MPMI 8(4): 484-91 (1995)); *Erwinia carotovora* (Cui, et. al., "The RsmA⁻ Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrp*N$_{Ecc}$ and Elicit a Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI 9(7): 565-73 (1966)); *Erwinia*

25   *stewartii* (Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microb. Inter. July 14-19, 1996 and Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," Ann. Mtg. Am. Phytopath. Soc. July 27-31, 1996); and *Pseudomonas syringae* pv. *syringae* (WO 94/26782 to Cornell Research Foundation, Inc.).

30   The present invention seeks to identify fragments of hypersensitive response elicitor proteins or polypeptides, which fragments elicit a hypersensitive response, and uses of such fragments.

SUMMARY OF THE INVENTION

The present invention is directed to an isolated fragment of an *Erwinia*
5   hypersensitive response elicitor protein or polypeptide where the fragment elicits a
hypersensitive response in plants. Also disclosed are isolated DNA molecules which
encode such fragments.

The fragments of hypersensitive response elicitors can be used to impart
disease resistance to plants, to enhance plant growth, and/or to control insects. This
10   involves applying the fragments in a non-infectious form to plants or plant seeds
under conditions effective to impart disease resistance, to enhance plant growth,
and/or to control insects on plants or plants grown from the plant seeds.

As an alternative to applying the fragments to plants or plant seeds in order to
impart disease resistance, to enhance plant growth, and/or to control insects on plants,
15   transgenic plants or plant seeds can be utilized. When utilizing transgenic plants, this
involves providing a transgenic plant transformed with a DNA molecule encoding a
fragment of a hypersensitive response elicitor protein or polypeptide which fragments
elicit a hypersensitive response in plants and growing the plant under conditions
effective to impart disease resistance, to enhance plant growth, and/or to control
20   insects in the plants or plants grown from the plant seeds. Alternatively, a transgenic
plant seed transformed with the DNA molecule encoding such a fragment can be
provided and planted in soil. A plant is then propagated under conditions effective to
impart disease resistance, to enhance plant growth, and/or to control insects on plants
or plants grown from the plant seeds.

25

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a deletion and proteolysis analysis for the *Erwinia amylovora*
hypersensitive response elicitor (i.e. harpin). A is the name of the harpin fragment. B
30   is the length of the fragment in amino acid residues. C indicates whether detectable
protein is produced. D states whether there is hypersensitive response (i.e., HR)
eliciting activity. The solid line indicates that there are additional amino acids which
are not harpin encoded, while the dashed line indicates the portion of the harpin that is

- 5 -

deleted. The numbers above the fragments in the box represent the amino acid residue present at the end of a given fragment; residue #1 is the N-terminus, and residue #403 is the C-terminus.

Figure 2 is a Western blot illustrating specific secretion of harpin$_{Ea}$, but not

5   harpin$_{Ea}$C31. Lane A, Ea273(pGP1-2) CFEP; lane B, Ea273(pGP1-2)(pCPP1104) CFEP; lane C, *E. coli* DH5α (pCPP1107) CFEP harpin size standard; lane D, BioRad low range molecular weight markers; lane E, Ea273(pGP1-2) supernatant; lane F, Ea273(pGP1 2)(pCPP1104) supernatant. The blot was probed with an anti-harpin$_{Ea}$ polyclonal antibody.

10   Figure 3 is an HR assay on tobacco leaf infiltrated as follows: (1) A, harpin$_{Ea}$ + raspberry IF; (2) B, harpin$_{Ea}$ + apple IF; (3) C, harpin$_{Ea}$ + tobacco IF; (4) D, harpin$_{Ea}$ + endoproteinase Glu-C; (5) E, harpin$_{Ea}$ + trypsin; (6) F, harpin$_{Ea}$; (7) G, tobacco IF; (8) H, endoproteinase Glu-C; (9) I, trypsin; and (10) J, harpin$_{Ea}$. IF refers to intracelluar fluids.

15   Figure 4 shows the digestion of harpin with endoproteinase Glu-C. Lane A is harpin; Lane B is harpin + endoproteinase Glu-C; Lane C is BioRad low range molecular weight markers.

Figure 5A shows the proteolysis of harpin. Coomassie blue stained polyacrylamide gel was loaded as follows: A, BioRad low range molecular weight

20   markers; B, IF-apple; C, IF-raspberry; D, IF-tobacco; E, harpin$_{Ea}$; F, harpin$_{Ea}$ + IF-apple; G, harpin$_{Ea}$ + IF-raspberry; H, harpin$_{Ea}$ + IF-tobacco.

Figure 5B shows a Coomassie Blue stained polyacrylamide gel loaded as follows: A, IF-tobacco; B, IF-tobacco + harpin$_{Ea}$; C, harpin$_{Ea}$; D, BioRad low range molecular weight markers; E, IF-tobacco + harpin$_{Ea}$ + PMSF. HR-eliciting activity of

25   the sample following proteolysis is denoted below the gel.

Figure 5C depicts whether proteolytic activity is present in IF from all plants tested. Intercellular fluid harvested from several plants was analyzed by PAGE in a gel containing 0.1% copolymerized gelatin. After washing to remove SDS and incubation to allow proteolysis of gelatin, the gels were stained to demonstrate the

30   presence of gelatinolytic activity. A, IF-apple; B, IF-tobacco; C, IF-cotoneaster; D, BioRad mw; E, endoproteinase Glu-C; and F, ground leaf extract-tobacco.

- 6 -

Figure 6 shows the refraction of elicitor-active peptides following proteolysis of harpin_{Ea} by tobacco IF. Absorbance was measured at 210 nm. Peak 1 contains peptides P91 and P95; peak 2 contains peptides P65 and P69.

Figure 7 shows the predicted proteolytic cleavage sites within harpin of

5   several tested proteinases, and the effect of these cleavages on activity of active harpin fragments. Residues potentially important for HR-eliciting activity, based on the loss of activity following further cleavage, are indicated by upward-pointing arrows at bottom.

Figure 8 shows the similarities near N-termini among harpins of *Erwinia* spp.

10   Underlined residues are present (identical or similar) in at least four out of the five proteins examined. Nine out of the first 26 residues are conserved in this manner.

Figures 9 A-B show Kyte-Doolittle hydropathy plots of bacterial HR-eliciting proteins. Ea, *E. amylovora* EA321; Est, *E. stewartii* DC283; Ech, *E. chrysanthemi* AC4150; Ecc, *E. cartovora* subsp. *carotovora*; Rs, *R, solanacearum*; Pss, *P. syringae*

15   pv. *syringae.*

Figure 10 shows truncated proteins of the hypersensitive response elicitor protein or polypeptide.

Figure 11 shows a list of synthesized oligonucleotide primers for construction of truncated harpin proteins. N represents the N-terminus (5' region), and C

20   represents the C-terminus (3' region). The primers correspond to the indicated sequence identification numbers for the present application: N1 (SEQ. ID. No. 1), N76 (SEQ. ID. No. 2), N99 (SEQ. ID. No. 3), N105 (SEQ. ID. No. 4), N110 (SEQ. ID. No. 5), N137 (SEQ. ID. No. 6), N150 (SEQ. ID. No. 7), N169 (SEQ. ID. No. 8), N210 (SEQ. ID. No. 9), N267 (SEQ. ID. No. 10), N343 (SEQ. ID. No. 11), C75

25   (SEQ. ID. No. 12), C104 (SEQ. ID. No. 13), C168 (SEQ. ID. No. 14), C180 (SEQ. ID. No. 15), C204 (SEQ. ID. No. 16), C209 (SEQ. ID. No. 17), C266 (SEQ. ID. No. 18), C342 (SEQ. ID. No. 19), and C403 (SEQ. ID. No. 20).

DETAILED DESCRIPTION OF THE INVENTION

30

The present invention is directed to isolated fragments of a hypersensitive response elicitor protein or polypeptide where the fragments elicit a hypersensitive response in plants. Also disclosed are DNA molecules encoding such fragments as

- 7 -

well as expression systems, host cells, and plants containing such molecules. Uses of the fragments themselves and the DNA molecules encoding them are disclosed.

The fragments of hypersensitive response elicitor polypeptides or proteins according to the present invention are derived from hypersensitive response elicitor polypeptides or proteins of a wide variety of fungal and bacterial pathogens. Such polypeptides or proteins are able to elicit local necrosis in plant tissue contacted by the elicitor. Examples of suitable bacterial sources of polypeptide or protein elicitors include *Erwinia*, *Pseudomonas*, and *Xanthamonas* species (e.g., the following bacteria: *Erwinia amylovora*, *Erwinia chrysanthemi*, *Erwinia stewartii*, *Erwinia carotovora*, *Pseudomonas syringae*, *Pseudomonas solancearum*, *Xanthomonas campestris*, and mixtures thereof).

An example of a fungal source of a hypersensitive response elicitor protein or polypeptide is *Phytophthora*. Suitable species of *Phytophthora* include *Phytophthora parasitica*, *Phytophthora cryptogea*, *Phytophthora cinnamomi*, *Phytophthora capsici*, *Phytophthora megasperma*, and *Phytophthora citrophthora*.

The hypersensitive response elicitor polypeptide or protein from *Erwinia chrysanthemi* has an amino acid sequence corresponding to SEQ. ID. No. 21 as follows:

```
Met Gln Ile Thr Ile Lys Ala His Ile Gly Gly Asp Leu Gly Val Ser
1           5           10              15

Gly Leu Gly Ala Gln Gly Leu Lys Gly Leu Asn Ser Ala Ala Ser Ser
        20              25              30

Leu Gly Ser Ser Val Asp Lys Leu Ser Ser Thr Ile Asp Lys Leu Thr
        35              40              45

Ser Ala Leu Thr Ser Met Met Phe Gly Gly Ala Leu Ala Gln Gly Leu
    50              55              60

Gly Ala Ser Ser Lys Gly Leu Gly Met Ser Asn Gln Leu Gly Gln Ser
65              70              75              80

Phe Gly Asn Gly Ala Gln Gly Ala Ser Asn Leu Leu Ser Val Pro Lys
            85              90              95

Ser Gly Gly Asp Ala Leu Ser Lys Met Phe Asp Lys Ala Leu Asp Asp
            100             105             110

Leu Leu Gly His Asp Thr Val Thr Lys Leu Thr Asn Gln Ser Asn Gln
            115             120             125
```

- 8 -

```
         Leu Ala Asn Ser Met Leu Asn Ala Ser Gln Met Thr Gln Gly Asn Met
             130                 135                 140
         Asn Ala Phe Gly Ser Gly Val Asn Asn Ala Leu Ser Ser Ile Leu Gly
             145                 150                 155                 160
5        Asn Gly Leu Gly Gln Ser Met Ser Gly Phe Ser Gln Pro Ser Leu Gly
                         165                 170                 175
         Ala Gly Gly Leu Gln Gly Leu Ser Gly Ala Gly Ala Phe Asn Gln Leu
                 180                 185                 190
10       Gly Asn Ala Ile Gly Met Gly Val Gly Gln Asn Ala Ala Leu Ser Ala
                 195                 200                 205
         Leu Ser Asn Val Ser Thr His Val Asp Gly Asn Asn Arg His Phe Val
             210                 215                 220
         Asp Lys Glu Asp Arg Gly Met Ala Lys Glu Ile Gly Gln Phe Met Asp
         225                 230                 235                 240
15       Gln Tyr Pro Glu Ile Phe Gly Lys Pro Glu Tyr Gln Lys Asp Gly Trp
                         245                 250                 255
         Ser Ser Pro Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser Lys
                 260                 265                 270
20       Pro Asp Asp Asp Gly Met Thr Gly Ala Ser Met Asp Lys Phe Arg Gln
                 275                 280                 285
         Ala Met Gly Met Ile Lys Ser Ala Val Ala Gly Asp Thr Gly Asn Thr
             290                 295                 300
         Asn Leu Asn Leu Arg Gly Ala Gly Gly Ala Ser Leu Gly Ile Asp Ala
         305                 310                 315                 320
25       Ala Val Val Gly Asp Lys Ile Ala Asn Met Ser Leu Gly Lys Leu Ala
                         325                 330                 335
         Asn Ala
```

30   This hypersensitive response elicitor polypeptide or protein has a molecular weight of 34 kDa, is heat stable, has a glycine content of greater than 16%, and contains substantially no cysteine. The *Erwinia chrysanthemi* hypersensitive response elicitor polypeptide or protein is encoded by a DNA molecule having a nucleotide sequence corresponding to SEQ. ID. No. 22 as follows:

35

```
CGATTTTACC CGGGTGAACG TGCTATGACC GACAGCATCA CGGTATTCGA CACCGTTACG    60
GCGTTTATGG CCGCGATGAA CCGGCATCAG GCGGCGCGCT GGTCGCCGCA ATCCGGCGTC    120
```

```
GATCTGGTAT TTCAGTTTGG GGACACCGGG CGTGAACTCA TGATGCAGAT TCAGCCGGGG      180

CAGCAATATC CCGGCATGTT GCGCACGCTG CTCGCTCGTC GTTATCAGCA GGCGGCAGAG      240

TGCGATGGCT GCCATCTGTG CCTGAACGGC AGCGATGTAT TGATCCTCTG GTGGCCGCTG      300

CCGTCGGATC CCGGCAGTTA TCCGCAGGTG ATCGAACGTT TGTTTGAACT GGCGGGAATG      360

ACGTTGCCGT CGCTATCCAT AGCACCGACG GCGCGTCCGC AGACAGGGAA CGGACGCGCC      420

CGATCATTAA GATAAAGGCG GCTTTTTTTA TTGCAAAACG GTAACGGTGA GGAACCGTTT      480

CACCGTCGGC GTCACTCAGT AACAAGTATC CATCATGATG CCTACATCGG GATCGGCGTG      540

GGCATCCGTT GCAGATACTT TTGCGAACAC CTGACATGAA TGAGGAAACG AAATTATGCA      600

AATTACGATC AAAGCGCACA TCGGCGGTGA TTTGGGCGTC TCCGGTCTGG GGCTGGGTGC      660

TCAGGGACTG AAAGGACTGA ATTCCGCGGC TTCATCGCTG GGTTCCAGCG TGGATAAACT      720

GAGCAGCACC ATCGATAAGT TGACCTCCGC GCTGACTTCG ATGATGTTTG GCGGCGCGCT      780

GGCGCAGGGG CTGGGCGCCA GCTCGAAGGG GCTGGGGATG AGCAATCAAC TGGGCCAGTC      840

TTTCGGCAAT GGCGCGCAGG GTGCGAGCAA CCTGCTATCC GTACCGAAAT CCGGCGGCGA      900

TGCGTTGTCA AAAATGTTTG ATAAAGCGCT GGACGATCTG CTGGGTCATG ACACCGTGAC      960

CAAGCTGACT AACCAGAGCA ACCAACTGGC TAATTCAATG CTGAACGCCA GCCAGATGAC     1020

CCAGGGTAAT ATGAATGCGT TCGGCAGCGG TGTGAACAAC GCACTGTCGT CCATTCTCGG     1080

CAACGGTCTC GGCCAGTCGA TGAGTGGCTT CTCTCAGCCT TCTCTGGGGG CAGGCGGCTT     1140

GCAGGGCCTG AGCGGCGCGG GTGCATTCAA CCAGTTGGGT AATGCCATCG GCATGGGCGT     1200

GGGGCAGAAT GCTGCGCTGA GTGCGTTGAG TAACGTCAGC ACCCACGTAG ACGGTAACAA     1260

CCGCCACTTT GTAGATAAAG AAGATCGCGG CATGGCGAAA GAGATCGGCC AGTTTATGGA     1320

TCAGTATCCG GAAATATTCG GTAAACCGGA ATACCAGAAA GATGGCTGGA GTTCGCCGAA     1380

GACGGACGAC AAATCCTGGG CTAAAGCGCT GAGTAAACCG GATGATGACG GTATGACCGG     1440

CGCCAGCATG GACAAATTCC GTCAGGCGAT GGGTATGATC AAAAGCGCGG TGGCGGGTGA     1500

TACCGGCAAT ACCAACCTGA ACCTGCGTGG CGCGGGCGGT GCATCGCTGG GTATCGATGC     1560

GGCTGTCGTC GGCGATAAAA TAGCCAACAT GTCGCTGGGT AAGCTGGCCA ACGCCTGATA     1620

ATCTGTGCTG GCCTGATAAA GCGGAAACGA AAAAGAGAC GGGGAAGCCT GTCTCTTTTC      1680

TTATTATGCG GTTTATGCGG TTACCTGGAC CGGTTAATCA TCGTCATCGA TCTGGTACAA     1740

ACGCACATTT TCCCGTTCAT TCGCGTCGTT ACGCGCCACA ATCGCGATGG CATCTTCCTC     1800

GTCGCTCAGA TTGCGCGGCT GATGGGGAAC GCCGGGTGGA ATATAGAGAA ACTCGCCGGC     1860
```

- 10 -

```
CAGATGGAGA CACGTCTGCG ATAAATCTGT GCCGTAACGT GTTTCTATCC GCCCCTTTAG    1920
CAGATAGATT GCGGTTTCGT AATCAACATG GTAATGCGGT TCCGCCTGTG CGCCGGCCGG    1980
GATCACCACA ATATTCATAG AAAGCTGTCT TGCACCTACC GTATCGCGGG AGATACCGAC    2040
AAAATAGGGC AGTTTTTGCG TGGTATCCGT GGGGTGTTCC GGCCTGACAA TCTTGAGTTG    2100
GTTCGTCATC ATCTTTCTCC ATCTGGGCGA CCTGATCGGT T                        2141
```

The hypersensitive response elicitor polypeptide or protein derived from *Erwinia amylovora* has an amino acid sequence corresponding to SEQ. ID. No. 23 as follows:

```
Met Ser Leu Asn Thr Ser Gly Leu Gly Ala Ser Thr Met Gln Ile Ser
1               5                   10                  15

Ile Gly Gly Ala Gly Gly Asn Asn Gly Leu Leu Gly Thr Ser Arg Gln
            20                  25                  30

Asn Ala Gly Leu Gly Gly Asn Ser Ala Leu Gly Leu Gly Gly Gly Asn
            35                  40                  45

Gln Asn Asp Thr Val Asn Gln Leu Ala Gly Leu Leu Thr Gly Met Met
    50                  55                  60

Met Met Met Ser Met Met Gly Gly Gly Gly Leu Met Gly Gly Gly Leu
65                  70                  75                  80

Gly Gly Gly Leu Gly Asn Gly Leu Gly Gly Ser Gly Gly Leu Gly Glu
            85                  90                  95

Gly Leu Ser Asn Ala Leu Asn Asp Met Leu Gly Gly Ser Leu Asn Thr
            100                 105                 110

Leu Gly Ser Lys Gly Gly Asn Asn Thr Thr Ser Thr Thr Asn Ser Pro
        115                 120                 125

Leu Asp Gln Ala Leu Gly Ile Asn Ser Thr Ser Gln Asn Asp Asp Ser
        130                 135                 140

Thr Ser Gly Thr Asp Ser Thr Ser Asp Ser Ser Asp Pro Met Gln Gln
145                 150                 155                 160

Leu Leu Lys Met Phe Ser Glu Ile Met Gln Ser Leu Phe Gly Asp Gly
                165                 170                 175

Gln Asp Gly Thr Gln Gly Ser Ser Ser Gly Gly Lys Gln Pro Thr Glu
            180                 185                 190

Gly Glu Gln Asn Ala Tyr Lys Lys Gly Val Thr Asp Ala Leu Ser Gly
        195                 200                 205
```

- 11 -

```
        Leu Met Gly Asn Gly Leu Ser Gln Leu Leu Gly Asn Gly Gly Leu Gly
            210                 215                 220

        Gly Gly Gln Gly Gly Asn Ala Gly Thr Gly Leu Asp Gly Ser Ser Leu
            225                 230                 235                 240

5       Gly Gly Lys Gly Leu Gln Asn Leu Ser Gly Pro Val Asp Tyr Gln Gln
                            245                 250                 255

        Leu Gly Asn Ala Val Gly Thr Gly Ile Gly Met Lys Ala Gly Ile Gln
                260                 265                 270

10      Ala Leu Asn Asp Ile Gly Thr His Arg His Ser Ser Thr Arg Ser Phe
            275                 280                 285

        Val Asn Lys Gly Asp Arg Ala Met Ala Lys Glu Ile Gly Gln Phe Met
            290                 295                 300

        Asp Gln Tyr Pro Glu Val Phe Gly Lys Pro Gln Tyr Gln Lys Gly Pro
        305                 310                 315                 320

15      Gly Gln Glu Val Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser
                        325                 330                 335

        Lys Pro Asp Asp Asp Gly Met Thr Pro Ala Ser Met Glu Gln Phe Asn
                340                 345                 350

20      Lys Ala Lys Gly Met Ile Lys Arg Pro Met Ala Gly Asp Thr Gly Asn
                355                 360                 365

        Gly Asn Leu Gln Ala Arg Gly Ala Gly Gly Ser Ser Leu Gly Ile Asp
            370                 375                 380

        Ala Met Met Ala Gly Asp Ala Ile Asn Asn Met Ala Leu Gly Lys Leu
        385                 390                 395                 400

25      Gly Ala Ala
```

This hypersensitive response elicitor polypeptide or protein has a molecular weight of about 39 kDa, has a pI of approximately 4.3, and is heat stable at 100°C for at least 10 minutes. This hypersensitive response elicitor polypeptide or protein has substantially no cysteine. The hypersensitive response elicitor polypeptide or protein derived from *Erwinia amylovora* is more fully described in Wei, Z.-M., R. J. Laby, C. H. Zumoff, D. W. Bauer, S.-Y. He, A. Collmer, and S. V. Beer, "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science 257:85-88 (1992), which is hereby incorporated by reference. The DNA

molecule encoding this polypeptide or protein has a nucleotide sequence corresponding to SEQ. ID. No. 24 as follows:

```
        AAGCTTCGGC ATGGCACGTT TGACCGTTGG GTCGGCAGGG TACGTTTGAA TTATTCATAA      60
   5    GAGGAATACG TTATGAGTCT GAATACAAGT GGGCTGGGAG CGTCAACGAT GCAAATTTCT     120
        ATCGGCGGTG CGGGCGGAAA TAACGGGTTG CTGGGTACCA GTCGCCAGAA TGCTGGGTTG     180
        GGTGGCAATT CTGCACTGGG GCTGGGCGGC GGTAATCAAA ATGATACCGT CAATCAGCTG     240
        GCTGGCTTAC TCACCGGCAT GATGATGATG ATGAGCATGA TGGGCGGTGG TGGGCTGATG     300
        GGCGGTGGCT TAGGCGGTGG CTTAGGTAAT GGCTTGGGTG GCTCAGGTGG CCTGGGCGAA     360
  10    GGACTGTCGA ACGCGCTGAA CGATATGTTA GGCGGTTCGC TGAACACGCT GGGCTCGAAA     420
        GGCGGCAACA ATACCACTTC AACAACAAAT TCCCCGCTGG ACCAGGCGCT GGGTATTAAC     480
        TCAACGTCCC AAAACGACGA TTCCACCTCC GGCACAGATT CCACCTCAGA CTCCAGCGAC     540
        CCGATGCAGC AGCTGCTGAA GATGTTCAGC GAGATAATGC AAAGCCTGTT TGGTGATGGG     600
        CAAGATGGCA CCCAGGGCAG TTCCTCTGGG GGCAAGCAGC CGACCGAAGG CGAGCAGAAC     660
  15    GCCTATAAAA AAGGAGTCAC TGATGCGCTG TCGGGCCTGA TGGGTAATGG TCTGAGCCAG     720
        CTCCTTGGCA ACGGGGGACT GGGAGGTGGT CAGGGCGGTA ATGCTGGCAC GGGTCTTGAC     780
        GGTTCGTCGC TGGGCGGCAA AGGGCTGCAA AACCTGAGCG GCCGGTGGA CTACCAGCAG     840
        TTAGGTAACG CCGTGGGTAC CGGTATCGGT ATGAAAGCGG GCATTCAGGC GCTGAATGAT     900
        ATCGGTACGC ACAGGCACAG TTCAACCCGT TCTTTCGTCA ATAAAGGCGA TCGGGCGATG     960
  20    GCGAAGGAAA TCGGTCAGTT CATGGACCAG TATCCTGAGG TGTTTGGCAA GCCGCAGTAC    1020
        CAGAAAGGCC CGGGTCAGGA GGTGAAAACC GATGACAAAT CATGGGCAAA AGCACTGAGC    1080
        AAGCCAGATG ACGACGGAAT GACACCAGCC AGTATGGAGC AGTTCAACAA AGCCAAGGGC    1140
        ATGATCAAAA GGCCCATGGC GGGTGATACC GGCAACGGCA ACCTGCAGGC ACGCGGTGCC    1200
        GGTGGTTCTT CGCTGGGTAT TGATGCCATG ATGGCCGGTG ATGCCATTAA CAATATGGCA    1260
  25    CTTGGCAAGC TGGGCGCGGC TTAAGCTT                                        1288
```

The hypersensitive response elicitor polypeptide or protein derived from *Pseudomonas syringae* has an amino acid sequence corresponding to SEQ. ID. No. 25

   30   as follows:

- 13 -

```
Met Gln Ser Leu Ser Leu Asn Ser Ser Ser Leu Gln Thr Pro Ala Met
1               5               10              15

        Ala Leu Val Leu Val Arg Pro Glu Ala Glu Thr Thr Gly Ser Thr Ser
                    20              25              30

5       Ser Lys Ala Leu Gln Glu Val Val Val Lys Leu Ala Glu Glu Leu Met
                35              40              45

        Arg Asn Gly Gln Leu Asp Asp Ser Ser Pro Leu Gly Lys Leu Leu Ala
            50              55              60

10      Lys Ser Met Ala Ala Asp Gly Lys Ala Gly Gly Ile Glu Asp Val
        65              70              75              80

        Ile Ala Ala Leu Asp Lys Leu Ile His Glu Lys Leu Gly Asp Asn Phe
                    85              90              95

        Gly Ala Ser Ala Asp Ser Ala Ser Gly Thr Gly Gln Gln Asp Leu Met
                    100             105             110

15      Thr Gln Val Leu Asn Gly Leu Ala Lys Ser Met Leu Asp Asp Leu Leu
                115             120             125

        Thr Lys Gln Asp Gly Gly Thr Ser Phe Ser Glu Asp Asp Met Pro Met
            130             135             140

        Leu Asn Lys Ile Ala Gln Phe Met Asp Asp Asn Pro Ala Gln Phe Pro
20      145             150             155             160

        Lys Pro Asp Ser Gly Ser Trp Val Asn Glu Leu Lys Glu Asp Asn Phe
                    165             170             175

        Leu Asp Gly Asp Glu Thr Ala Ala Phe Arg Ser Ala Leu Asp Ile Ile
                    180             185             190

25      Gly Gln Gln Leu Gly Asn Gln Gln Ser Asp Ala Gly Ser Leu Ala Gly
                    195             200             205

        Thr Gly Gly Gly Leu Gly Thr Pro Ser Ser Phe Ser Asn Asn Ser Ser
            210             215             220

        Val Met Gly Asp Pro Leu Ile Asp Ala Asn Thr Gly Pro Gly Asp Ser
30      225             230             235             240

        Gly Asn Thr Arg Gly Glu Ala Gly Gln Leu Ile Gly Glu Leu Ile Asp
                    245             250             255

        Arg Gly Leu Gln Ser Val Leu Ala Gly Gly Gly Leu Gly Thr Pro Val
                    260             265             270

35      Asn Thr Pro Gln Thr Gly Thr Ser Ala Asn Gly Gly Gln Ser Ala Gln
                    275             280             285
```

- 14 -

```
        Asp Leu Asp Gln Leu Leu Gly Gly Leu Leu Leu Lys Gly Leu Glu Ala
            290                 295                 300

        Thr Leu Lys Asp Ala Gly Gln Thr Gly Thr Asp Val Gln Ser Ser Ala
        305                 310                 315                 320
  5     Ala Gln Ile Ala Thr Leu Leu Val Ser Thr Leu Leu Gln Gly Thr Arg
                        325                 330                 335

        Asn Gln Ala Ala Ala
                    340
```

10

This hypersensitive response elicitor polypeptide or protein has a molecular weight of 34-35 kDa. It is rich in glycine (about 13.5%) and lacks cysteine and tyrosine. Further information about the hypersensitive response elicitor derived from *Pseudomonas syringae* is found in He, S. Y., H. C. Huang, and A. Collmer,

15 "*Pseudomonas syringae* pv. *syringae* Harpin$_{Pss}$: a Protein that is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference. The DNA molecule encoding the hypersensitive response elicitor from *Pseudomonas syringae* has a nucleotide sequence corresponding to SEQ. ID. No. 26 as follows:

20

```
        ATGCAGAGTC TCAGTCTTAA CAGCAGCTCG CTGCAAACCC CGGCAATGGC CCTTGTCCTG      60

        GTACGTCCTG AAGCCGAGAC GACTGGCAGT ACGTCGAGCA AGGCGCTTCA GGAAGTTGTC     120

        GTGAAGCTGG CCGAGGAACT GATGCGCAAT GGTCAACTCG ACGACAGCTC GCCATTGGGA     180

        AAACTGTTGG CCAAGTCGAT GGCCGCAGAT GGCAAGGCGG GCGGCGGTAT TGAGGATGTC     240

  25    ATCGCTGCGC TGGACAAGCT GATCCATGAA AAGCTCGGTG ACAACTTCGG CGCGTCTGCG     300

        GACAGCGCCT CGGGTACCGG ACAGCAGGAC CTGATGACTC AGGTGCTCAA TGGCCTGGCC     360

        AAGTCGATGC TCGATGATCT TCTGACCAAG CAGGATGGCG GGACAAGCTT CTCCGAAGAC     420

        GATATGCCGA TGCTGAACAA GATCGCGCAG TTCATGGATG ACAATCCCGC ACAGTTTCCC     480

        AAGCCGGACT CGGGCTCCTG GGTGAACGAA CTCAAGGAAG ACAACTTCCT TGATGGCGAC     540

  30    GAAACGGCTG CGTTCCGTTC GGCACTCGAC ATCATTGGCC AGCAACTGGG TAATCAGCAG     600

        AGTGACGCTG GCAGTCTGGC AGGGACGGGT GGAGGTCTGG GCACTCCGAG CAGTTTTTCC     660

        AACAACTCGT CCGTGATGGG TGATCCGCTG ATCGACGCCA ATACCGGTCC CGGTGACAGC     720

        GGCAATACCC GTGGTGAAGC GGGGCAACTG ATCGGCGAGC TTATCGACCG TGGCCTGCAA     780

        TCGGTATTGG CCGGTGGTGG ACTGGGCACA CCCGTAAACA CCCCGCAGAC CGGTACGTCG     840
```

- 15 -

```
GCGAATGGCG GACAGTCCGC TCAGGATCTT GATCAGTTGC TGGGCGGCTT GCTGCTCAAG        900

GGCCTGGAGG CAACGCTCAA GGATGCCGGG CAAACAGGCA CCGACGTGCA GTCGAGCGCT        960

GCGCAAATCG CCACCTTGCT GGTCAGTACG CTGCTGCAAG GCACCCGCAA TCAGGCTGCA       1020

GCCTGA                                                                  1026
```

The hypersensitive response elicitor polypeptide or protein derived from *Pseudomonas solanacearum* has an amino acid sequence corresponding to SEQ. ID. No. 27 as follows:

```
Met Ser Val Gly Asn Ile Gln Ser Pro Ser Asn Leu Pro Gly Leu Gln
1            5               10               15

Asn Leu Asn Leu Asn Thr Asn Thr Asn Ser Gln Gln Ser Gly Gln Ser
         20               25               30

Val Gln Asp Leu Ile Lys Gln Val Glu Lys Asp Ile Leu Asn Ile Ile
         35               40               45

Ala Ala Leu Val Gln Lys Ala Ala Gln Ser Ala Gly Gly Asn Thr Gly
     50               55               60

Asn Thr Gly Asn Ala Pro Ala Lys Asp Gly Asn Ala Asn Ala Gly Ala
65               70               75               80

Asn Asp Pro Ser Lys Asn Asp Pro Ser Lys Ser Gln Ala Pro Gln Ser
             85               90               95

Ala Asn Lys Thr Gly Asn Val Asp Asp Ala Asn Asn Gln Asp Pro Met
         100              105              110

Gln Ala Leu Met Gln Leu Leu Glu Asp Leu Val Lys Leu Leu Lys Ala
         115              120              125

Ala Leu His Met Gln Gln Pro Gly Gly Asn Asp Lys Gly Asn Gly Val
         130              135              140

Gly Gly Ala Asn Gly Ala Lys Gly Ala Gly Gly Gln Gly Gly Leu Ala
145              150              155              160

Glu Ala Leu Gln Glu Ile Glu Gln Ile Leu Ala Gln Leu Gly Gly Gly
             165              170              175

Gly Ala Gly Ala Gly Gly Ala Gly Gly Gly Val Gly Gly Ala Gly Gly
         180              185              190

Ala Asp Gly Gly Ser Gly Ala Gly Gly Ala Gly Gly Ala Asn Gly Ala
         195              200              205
```

- 16 -

```
Asp Gly Gly Asn Gly Val Asn Gly Asn Gln Ala Asn Gly Pro Gln Asn
    210             215             220
```

```
Ala Gly Asp Val Asn Gly Ala Asn Gly Ala Asp Asp Gly Ser Glu Asp
225             230             235             240
```

5
```
Gln Gly Gly Leu Thr Gly Val Leu Gln Lys Leu Met Lys Ile Leu Asn
            245             250             255
```

```
Ala Leu Val Gln Met Met Gln Gln Gly Gly Leu Gly Gly Gly Asn Gln
            260             265             270
```

10
```
Ala Gln Gly Gly Ser Lys Gly Ala Gly Asn Ala Ser Pro Ala Ser Gly
    275             280             285
```

```
Ala Asn Pro Gly Ala Asn Gln Pro Gly Ser Ala Asp Asp Gln Ser Ser
    290             295             300
```

```
Gly Gln Asn Asn Leu Gln Ser Gln Ile Met Asp Val Val Lys Glu Val
305             310             315             320
```

15
```
Val Gln Ile Leu Gln Gln Met Leu Ala Ala Gln Asn Gly Gly Ser Gln
            325             330             335
```

```
Gln Ser Thr Ser Thr Gln Pro Met
            340
```

It is encoded by a DNA molecule having a nucleotide sequence corresponding SEQ.
20 ID. No. 28 as follows:

```
ATGTCAGTCG GAAACATCCA GAGCCCGTCG AACCTCCCGG GTCTGCAGAA CCTGAACCTC     60

AACACCAACA CCAACAGCCA GCAATCGGGC CAGTCCGTGC AAGACCTGAT CAAGCAGGTC    120

GAGAAGGACA TCCTCAACAT CATCGCAGCC CTCGTGCAGA AGGCCGCACA GTCGGCGGGC    180

GGCAACACCG GTAACACCGG CAACGCGCCG GCGAAGGACG GCAATGCCAA CGCGGGCGCC    240
```
25
```
AACGACCCGA GCAAGAACGA CCCGAGCAAG AGCCAGGCTC GCAGTCGGC CAACAAGACC    300

GGCAACGTCG ACGACGCCAA CAACCAGGAT CCGATGCAAG CGCTGATGCA GCTGCTGGAA    360

GACCTGGTGA AGCTGCTGAA GGCGGCCCTG CACATGCAGC AGCCCGGCGG CAATGACAAG    420

GGCAACGGCG TGGGCGGTGC CAACGGCGCC AAGGGTGCCG GCGGCCAGGG CGGCCTGGCC    480

GAAGCGCTGC AGGAGATCGA GCAGATCCTC GCCCAGCTCG GCGGCGGCGG TGCTGGCGCC    540
```
30
```
GGCGGCGCGG GTGGCGGTGT CGGCGGTGCT GGTGGCGCGG ATGGCGGCTC CGGTGCGGGT    600

GGCGCAGGCG GTGCGAACGG CGCCGACGGC GGCAATGGCG TGAACGGCAA CCAGGCGAAC    660

GGCCCGCAGA ACGCAGGCGA TGTCAACGGT GCCAACGGCG CGGATGACGG CAGCGAAGAC    720

CAGGGCGGCC TCACCGGCGT GCTGCAAAAG CTGATGAAGA TCCTGAACGC GCTGGTGCAG    780
```

```
ATGATGCAGC AAGGCGGCCT CGGCGGCGGC AACCAGGCGC AGGGCGGCTC GAAGGGTGCC    840

GGCAACGCCT CGCCGGCTTC CGGCGCGAAC CCGGGCGCGA ACCAGCCCGG TTCGGCGGAT    900

GATCAATCGT CCGGCCAGAA CAATCTGCAA TCCCAGATCA TGGATGTGGT GAAGGAGGTC    960

GTCCAGATCC TGCAGCAGAT GCTGGCGGCG CAGAACGGCG GCAGCCAGCA GTCCACCTCG   1020

ACGCAGCCGA TGTAA                                                   1035
```

Further information regarding the hypersensitive response elicitor polypeptide or protein derived from *Pseudomonas solanacearum* is set forth in Arlat, M., F. Van Gijsegem, J. C. Huet, J. C. Pemollet, and C. A. Boucher, "PopA1, a Protein which Induces a Hypersensitive-like Response in Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-533 (1994), which is hereby incorporated by reference.

The hypersensitive response elicitor polypeptide or protein from *Xanthomonas campestris* pv. glycines has an amino acid sequence corresponding to SEQ. ID. No. 29 as follows:

```
Thr Leu Ile Glu Leu Met Ile Val Val Ala Ile Ile Ala Ile Leu Ala
1           5               10              15

Ala Ile Ala Leu Pro Ala Tyr Gln Asp Tyr
            20              25
```

This sequence is an amino terminal sequence having only 26 residues from the hypersensitive response elicitor polypeptide or protein of *Xanthomonas campestris* pv. glycines. It matches with fimbrial subunit proteins determined in other *Xanthomonas campestris* pathovars.

The hypersensitive response elicitor polypeptide or protein from *Xanthomonas campestris pv. pelargonii* is heat stable, protease sensitive, and has a molecular weight of 20 kDa. It includes an amino acid sequence corresponding to SEQ. ID. No. 30 as follows:

```
Ser Ser Gln Gln Ser Pro Ser Ala Gly Ser Glu Gln Gln Leu Asp Gln
1           5               10              15

Leu Leu Ala Met
            20
```

Isolation of *Erwinia carotovora* hypersensitive response elictor protein or polypeptide is described in Cui et al., "The RsmA Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrp* N$_{Ecc}$ and Elicit a Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI, 9(7):565-73 (1996), which is

5   hereby incorporated by reference. The hypersensitive response elicitor protein or polypeptide of *Erwinia stewartii* is set forth in Ahmad et al., "Harpin is Not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microbe Interact., July ___ Ahmad, et al., "Harpin is Not Necessary for the Pathogenicity of *Erwinia* ___ on Maize," Ann. Mtg. Am.

10   Phytopath. Soc., July 27-31, 1996, which are hereby incorporated by reference.

Hypersensitive response elicitor proteins or polypeptides from *Phytophthora parasitica, Phytophthora cryptogea, Phytophthora cinnamoni, Phytophthora capsici, Phytophthora megasperma,* and *Phytophora citrophthora* are described in Kaman, et al., "Extracellular Protein Elicitors from Phytophthora: Most Specificity and

15   Induction of Resistance to Bacterial and Fungal Phytopathogens," Molec. Plant-Microbe Interact., 6(1):15-25 (1993), Ricci et al., "Structure and Activity of Proteins from Pathogenic Fungi Phytophthora Eliciting Necrosis and Acquired Resistance in Tobacco," Eur. J. Biochem., 183:555-63 (1989), Ricci et al., "Differential Production of Parasiticein, and Elicitor of Necrosis and Resistance in Tobacco, by Isolates of

20   Phytophthora parasitica," Plant Path. 41:298-307 (1992), Baillreul et al, "A New Elicitor of the Hypersensitive Response in Tobacco: A Fungal Glycoprotein Elicits Cell Death, Expression of Defence Genes, Production of Salicylic Acid, and Induction of Systemic Acquired Resistance," Plant J., 8(4):551-60 (1995), and Bonnet et al., "Acquired Resistance Triggered by Elicitors in Tobacco and Other Plants," Eur. J.

25   Plant Path., 102:181-92 (1996), which are hereby incorporated by reference.

The above elicitors are exemplary. Other elicitors can be identified by growing fungi or bacteria that elicit a hypersensitive response under which genes encoding an elicitor are expressed. Cell-free preparations from culture supernatants can be tested for elicitor activity (i.e. local necrosis) by using them to infiltrate

30   appropriate plant tissues.

Fragments of the above hypersensitive response elicitor polypeptides or proteins as well as fragments of full length elicitors from other pathogens are encompassed by the method of the present invention.

Suitable fragments can be produced by several means. In the first, subclones
5   of the gene encoding a known elicitor protein are produced by conventional molecular genetic manipulation by subcloning gene fragments. The subclones then are expressed *in vitro* or *in vivo* in bacterial cells to yield a smaller protein or peptide that can be tested for elicitor activity according to the procedure described below.

As an alternative, fragments of an elicitor protein can be produced by
10   digestion of a full-length elicitor protein with proteolytic enzymes like chymotrypsin or *Staphylococcus* proteinase A, or trypsin. Different proteolytic enzymes are likely to cleave elicitor proteins at different sites based on the amino acid sequence of the elicitor protein. Some of the fragments that result from proteolysis may be active elicitors of resistance.

15   In another approach, based on knowledge of the primary structure of the protein, fragments of the elicitor protein gene may be synthesized by using the PCR technique together with specific sets of primers chosen to represent particular portions of the protein. These then would be cloned into an appropriate vector for expression of a truncated peptide or protein.

20   Chemical synthesis can also be used to make suitable fragments. Such a synthesis is carried out using known amino acid sequences for the elicitor being produced. Alternatively, subjecting a full length elicitor to high temperatures and pressures will produce fragments. These fragments can then be separated by conventional procedures (e.g., chromatography, SDS-PAGE).

25   An example of suitable fragments of an *Erwinia* hypersensitive response elicitor which fragments elicit a hypersensitive response are fragments of the *Erwinia amylovora* hypersensitive response elicitor. Suitable fragments include a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, or an internal fragment of the amino
30   acid sequence of SEQ. ID. No. 23. The C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 can span amino acids 105 and 403 of SEQ. ID. No. 23. The N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 can span the

following amino acids of SEQ. ID. No. 23: 1 and 98, 1 and 104, 1 and 122, 1 and 168, 1 and 218, 1 and 266, 1 and 342, 1 and 321, and 1 and 372. The internal fragment of the amino acid sequence of SEQ. ID. No. 23 can span the following amino acids of SEQ. ID. No. 23: 76 and 209, 105 and 209, 99 and 209, 137 and 204,

5   137 and 200, 109 and 204, 109 and 200, 137 and 180, and 105 and 180. Other suitable fragments can be identified in accordance with the present invention.

Variants may be made by, for example, the deletion or addition of amino acids that have minimal influence on the properties, secondary structure and hydropathic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal

10   (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification, or identification of the polypeptide.

The fragment of the present invention is preferably produced in

15   purified form (preferably at least about 60%, more preferably 80%, pure) by conventional techniques. Typically, the fragment of the present invention is produced but not secreted into the growth medium of recombinant host cells. Alternatively, the protein or polypeptide of the present invention is secreted into growth medium. In the case of unsecreted protein, to isolate the protein fragment, the host cell (e.g., *E. coli*)

20   carrying a recombinant plasmid is propagated, lysed by sonication, heat, or chemical treatment, and the homogenate is centrifuged to remove bacterial debris. The supernatant is then subjected to heat treatment and the fragment is separated by centrifugation. The supernatant fraction containing the fragment is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the

25   fragment. If necessary, the protein fraction may be further purified by ion exchange or HPLC.

The DNA molecule encoding the fragment of the hypersensitive response elicitor polypeptide or protein can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule

30   into an expression system to which the DNA molecule is heterologous (i.e. not normally present). The heterologous DNA molecule is inserted into the expression system or vector in proper sense orientation and correct reading frame. The vector

- 21 -

contains the necessary elements for the transcription and translation of the inserted protein-coding sequences.

U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of

5   recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including procaryotic organisms and eucaryotic cells grown in tissue culture.

Recombinant genes may also be introduced into viruses, such as vaccina virus.

10   Recombinant viruses can be generated by transection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC1084, pUC8, pUC9, pUC18, pUC19,

15   pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see F.W. Studier et. al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes," Gene Expression Technology vol. 185 (1990), which is hereby incorporated by

20   reference), and any derivatives thereof. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, mobilization, or electroporation. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York

25   (1989), which is hereby incorporated by reference.

A variety of host-vector systems may be utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA;

30   microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria. The expression

elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

5   Different genetic signals and processing events control many levels of gene expression (e.g., DNA transcription and messenger RNA (mRNA) translation).

Transcription of DNA is dependent upon the presence of a promotor which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eucaryotic promotors differ from those of procaryotic promotors. Furthermore, eucaryotic promotors and accompanying genetic

10   signals may not be recognized in or may not function in a procaryotic system, and, further, procaryotic promotors are not recognized and do not function in eucaryotic cells.

Similarly, translation of mRNA in procaryotes depends upon the presence of the proper procaryotic signals which differ from those of eucaryotes. Efficient

15   translation of mRNA in procaryotes requires a ribosome binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably

20   promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression, see Roberts and Lauer, Methods in Enzymology, 68:473 (1979), which is hereby incorporated by reference.

Promotors vary in their "strength" (i.e. their ability to promote transcription).

25   For the purposes of expressing a cloned gene, it is desirable to use strong promotors in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promotors may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promotors such as the T7 phage promoter, *lac* promotor, *trp* promotor, *rec*A

30   promotor, ribosomal RNA promotor, the P_R and P_L promotors of coliphage lambda and others, including but not limited, to *lac*UV5, *ompF*, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a

hybrid *trp-lac*UV5 (*tac*) promotor or other *E. coli* promotors produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

Bacterial host cell strains and expression vectors may be chosen which inhibit
5   the action of the promotor unless specifically induced. In certain operations, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

10   Specific initiation signals are also required for efficient gene transcription and translation in procaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promotor, may also contain any combination of various "strong" transcription and/or
15   translation initiation signals. For instance, efficient translation in *E. coli* requires an SD sequence about 7-9 bases 5' to the initiation codon ("ATG") to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or
20   from the *E. coli* tryptophan E, D, C, B or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

Once the isolated DNA molecule encoding the fragment of a hypersensitive response elicitor polypeptide or protein has been cloned into an expression system, it
25   is ready to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system. Suitable host cells include, but are not limited to, bacteria, virus, yeast, mammalian cells, insect, plant, and the like.

The present invention further relates to methods of imparting disease
30   resistance to plants, enhancing plant growth, and/or effecting insect control for plants. These methods involve applying the fragment of a hypersensitive response elicitor polypeptide or protein, which fragment itself elicits a hypersensitive response, in a

- 24 -

non-infectious form to all or part of a plant or a plant seed under conditions effective for the fragment to impart disease resistance, enhance growth, and/or control insects. Alternatively, these fragments of a hypersensitive response elicitor protein or polypeptide can be applied to plants such that seeds recovered from such plants

5   themselves are able to impart disease resistance in plants, to enhance plant growth, and/or to effect insect control.

As an alternative to applying a fragment of a hypersensitive response elicitor polypeptide or protein to plants or plant seeds in order to impart disease resistance in plants, to effect plant growth, and/or to control insects on the plants or plants grown

10   from the seeds, transgenic plants or plant seeds can be utilized. When utilizing transgenic plants, this involves providing a transgenic plant transformed with a DNA molecule encoding a fragment of a hypersensitive response elicitor polypeptide or protein, which fragment elicits a hypersensitive response, and growing the plant under conditions effective to permit that DNA molecule to impart disease resistance to

15   plants, to enhance plant growth, and/or to control insects. Alternatively, a transgenic plant seed transformed with a DNA molecule encoding a fragment of a hypersensitive response elicitor polypeptide or protein which fragment elicits a hypersensitive response can be provided and planted in soil. A plant is then propagated from the planted seed under conditions effective to permit that DNA molecule to impart

20   disease resistance to plants, to enhance plant growth, and/or to control insects.

The embodiment of the present invention where the hypersensitive response elicitor polypeptide or protein is applied to the plant or plant seed can be carried out in a number of ways, including: 1) application of an isolated fragment or 2) application of bacteria which do not cause disease and are transformed with a genes encoding the

25   fragment. In the latter embodiment, the fragment can be applied to plants or plant seeds by applying bacteria containing the DNA molecule encoding the fragment of the hypersensitive response elicitor polypeptide or protein which fragment elicits a hypersensitive response. Such bacteria must be capable of secreting or exporting the fragment so that the fragment can contact plant or plant seeds cells. In these

30   embodiments, the fragment is produced by the bacteria *in planta* or on seeds or just prior to introduction of the bacteria to the plants or plant seeds.

The methods of the present invention can be utilized to treat a wide variety of plants or their seeds to impart disease resistance, enhance growth, and/or control insects. Suitable plants include dicots and monocots. More particularly, useful crop plants can include: alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn,

5 potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane. Examples of suitable ornamental plants are: *Arabidopsis thaliana*,

10 *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

With regard to the use of the fragments of the hypersensitive response elicitor protein or polypeptide of the present invention in imparting disease resistance, absolute immunity against infection may not be conferred, but the severity of the disease is reduced and symptom development is delayed. Lesion number, lesion size,

15 and extent of sporulation of fungal pathogens are all decreased. This method of imparting disease resistance has the potential for treating previously untreatable diseases, treating diseases systemically which might not be treated separately due to cost, and avoiding the use of infectious agents or environmentally harmful materials.

The method of imparting pathogen resistance to plants in accordance with the

20 present invention is useful in imparting resistance to a wide variety of pathogens including viruses, bacteria, and fungi. Resistance, *inter alia*, to the following viruses can be achieved by the method of the present invention: *Tobacco mosaic virus* and *Tomato mosaic virus*. Resistance, *inter alia*, to the following bacteria can also be imparted to plants in accordance with present invention: *Pseudomonas solancearum*,

25 *Pseudomonas syringae* pv. *tabaci*, and *Xanthamonas campestris* pv. *pelargonii*. Plants can be made resistant, *inter alia*, to the following fungi by use of the method of the present invention: *Fusarium oxysporum* and *Phytophthora infestans*.

With regard to the use of the fragments of the hypersensitive response elicitor protein or polypeptide of the present invention to enhance plant growth, various forms

30 of plant growth enhancement or promotion can be achieved. This can occur as early as when plant growth begins from seeds or later in the life of a plant. For example, plant growth according to the present invention encompasses greater yield, increased

- 26 -

quantity of seeds produced, increased percentage of seeds germinated, increased plant size, greater biomass, more and bigger fruit, earlier fruit coloration, and earlier fruit and plant maturation. As a result, the present invention provides significant economic benefit to growers. For example, early germination and early maturation permit crops

5   to be grown in areas where short growing seasons would otherwise preclude their growth in that locale. Increased percentage of seed germination results in improved crop stands and more efficient seed use. Greater yield, increased size, and enhanced biomass production allow greater revenue generation from a given plot of land.

Another aspect of the present invention is directed to effecting any form of

10 insect control for plants. For example, insect control according to the present invention encompasses preventing insects from contacting plants to which the hypersensitive response elicitor has been applied, preventing direct insect damage to plants by feeding injury, causing insects to depart from such plants, killing insects proximate to such plants, interfering with insect larval feeding on such plants,

15 preventing insects from colonizing host plants, preventing colonizing insects from releasing phytotoxins, etc. The present invention also prevents subsequent disease damage to plants resulting from insect infection.

The present invention is effective against a wide variety of insects. European corn borer is a major pest of corn (dent and sweet corn) but also feeds on over 200

20 plant species including green, wax, and lima beans and edible soybeans, peppers, potato, and tomato plus many weed species. Additional insect larval feeding pests which damage a wide variety of vegetable crops include the following: beet armyworm, cabbage looper, corn ear worm, fall armyworm, diamondback moth, cabbage root maggot, onion maggot, seed corn maggot, pickleworm (melonworm),

25 pepper maggot, tomato pinworm, and maggots. Collectively, this group of insect pests represents the most economically important group of pests for vegetable production worldwide.

The method of the present invention involving application of the fragment of a hypersensitive response elicitor polypeptide or protein, which fragment elicits a

30 hypersensitive response, can be carried out through a variety of procedures when all or part of the plant is treated, including leaves, stems, roots, etc. This may (but need not) involve infiltration of the hypersensitive response elicitor polypeptide or protein

- 27 -

into the plant. Suitable application methods include high or low pressure spraying, injection, and leaf abrasion proximate to when elicitor application takes place. When treating plant seeds or propagules (e.g., cuttings), in accordance with the application embodiment of the present invention, the fragment of the hypersensitive response

5   elicitor protein or polypeptide, in accordance with present invention, can be applied by low or high pressure spraying, coating, immersion, or injection. Other suitable application procedures can be envisioned by those skilled in the art provided they are able to effect contact of the fragment with cells of the plant or plant seed. Once treated with the fragment of the hypersensitive response elicitor of the present

10   invention, the seeds can be planted in natural or artificial soil and cultivated using conventional procedures to produce plants. After plants have been propagated from seeds treated in accordance with the present invention, the plants may be treated with one or more applications of the fragment of the hypersensitive response elicitor protein or polypeptide or whole elicitors to impart disease resistance to plants, to

15   enhance plant growth, and/or to control insects on the plants.

The fragment of the hypersensitive response elicitor polypeptide or protein, in accordance with the present invention, can be applied to plants or plant seeds alone or in a mixture with other materials. Alternatively, the fragment can be applied separately to plants with other materials being applied at different times.

20   A composition suitable for treating plants or plant seeds in accordance with the application embodiment of the present invention contains a fragment of a hypersensitive response elicitor polypeptide or protein which fragment elicits a hypersensitive response in a carrier. Suitable carriers include water, aqueous solutions, slurries, or dry powders. In this embodiment, the composition contains

25   greater than 500 nM of the fragment.

Although not required, this composition may contain additional additives including fertilizer, insecticide, fungicide, nematacide, and mixtures thereof. Suitable fertilizers include $(NH_4)_2NO_3$. An example of a suitable insecticide is Malathion. Useful fungicides include Captan.

30   Other suitable additives include buffering agents, wetting agents, coating agents, and abrading agents. These materials can be used to facilitate the process of the present invention. In addition, the hypersensitive response eliciting fragment can

- 28 -

be applied to plant seeds with other conventional seed formulation and treatment materials, including clays and polysaccharides.

In the alternative embodiment of the present invention involving the use of transgenic plants and transgenic seeds, a hypersensitive response eliciting fragment

5 need not be applied topically to the plants or seeds. Instead, transgenic plants transformed with a DNA molecule encoding such a fragment are produced according to procedures well known in the art.

The vector described above can be microinjected directly into plant cells by use of micropipettes to transfer mechanically the recombinant DNA. Crossway, Mol.

10 Gen. Genetics, 202:179-85 (1985), which is hereby incorporated by reference. The genetic material may also be transferred into the plant cell using polyethylene glycol. Krens, et al., Nature, 296:72-74 (1982), which is hereby incorporated by reference.

Another approach to transforming plant cells with a gene which imparts resistance to pathogens is particle bombardment (also known as biolistic

15 transformation) of the host cell. This can be accomplished in one of several ways. The first involves propelling inert or biologically active particles at cells. This technique is disclosed in U.S. Patent Nos. 4,945,050, 5,036,006, and 5,100,792, all to Sanford et al., which are hereby incorporated by reference. Generally, this procedure involves propelling inert or biologically active particles at the cells under conditions

20 effective to penetrate the outer surface of the cell and to be incorporated within the interior thereof. When inert particles are utilized, the vector can be introduced into the cell by coating the particles with the vector containing the heterologous DNA. Alternatively, the target cell can be surrounded by the vector so that the vector is carried into the cell by the wake of the particle. Biologically active particles (e.g.,

25 dried bacterial cells containing the vector and heterologous DNA) can also be propelled into plant cells.

Yet another method of introduction is fusion of protoplasts with other entities, either minicells, cells, lysosomes, or other fusible lipid-surfaced bodies. Fraley, et al., Proc. Natl. Acad. Sci. USA, 79:1859-63 (1982), which is hereby incorporated by

30 reference.

The DNA molecule may also be introduced into the plant cells by electroporation. Fromm et al., Proc. Natl. Acad. Sci. USA, 82:5824 (1985), which is

- 29 -

hereby incorporated by reference. In this technique, plant protoplasts are electroporated in the presence of plasmids containing the expression cassette. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the

5   cell wall, divide, and regenerate.

Another method of introducing the DNA molecule into plant cells is to infect a plant cell with *Agrobacterium tumefaciens* or *A. rhizogenes* previously transformed with the gene. Under appropriate conditions known in the art, the transformed plant cells are grown to form shoots or roots, and develop further into plants. Generally,

10   this procedure involves inoculating the plant tissue with a suspension of bacteria and incubating the tissue for 48 to 72 hours on regeneration medium without antibiotics at 25-28°C.

Agrobacterium is a representative genus of the gram-negative family Rhizobiaceae. Its species are responsible for crown gall (*A. tumefaciens*) and hairy

15   root disease (*A. rhizogenes*). The plant cells in crown gall tumors and hairy roots are induced to produce amino acid derivatives known as opines, which are catabolized only by the bacteria. The bacterial genes responsible for expression of opines are a convenient source of control elements for chimeric expression cassettes. In addition, assaying for the presence of opines can be used to identify transformed tissue.

20   Heterologous genetic sequences can be introduced into appropriate plant cells, by means of the Ti plasmid of *A. tumefaciens* or the Ri plasmid of *A. rhizogenes*. The Ti or Ri plasmid is transmitted to plant cells on infection by Agrobacterium and is stably integrated into the plant genome. J. Schell, Science, 237:1176-83 (1987), which is hereby incorporated by reference.

25   After transformation, the transformed plant cells must be regenerated.

Plant regeneration from cultured protoplasts is described in Evans et al., Handbook of Plant Cell Cultures, Vol. 1: (MacMillan Publishing Co., New York, 1983); and Vasil I.R. (ed.), Cell Culture and Somatic Cell Genetics of Plants, Acad. Press, Orlando, Vol. I, 1984, and Vol. III (1986), which are hereby incorporated by

30   reference.

It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to, all major species of sugarcane, sugar beets, cotton, fruit trees, and legumes.

Means for regeneration vary from species to species of plants, but generally a

5   suspension of transformed protoplasts or a petri plate containing transformed explants is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced in the callus tissue. These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and

10   cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is usually reproducible and repeatable.

After the expression cassette is stably incorporated in transgenic plants, it can

15   be transferred to other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

Once transgenic plants of this type are produced, the plants themselves can be cultivated in accordance with conventional procedure with the presence of the gene encoding the hypersensitive response eliciting fragment resulting in disease

20   resistance, enhanced plant growth, and/or control of insects on the plant. Alternatively, transgenic seeds or propagules (e.g., cuttings) are recovered from the transgenic plants. The seeds can then be planted in the soil and cultivated using conventional procedures to produce transgenic plants. The transgenic plants are propagated from the planted transgenic seeds under conditions effective to impart

25   disease resistance to plants, to enhance plant growth, and/or to control insects. While not wishing to be bound by theory, such disease resistance, growth enhancement, and/or insect control may be RNA mediated or may result from expression of the polypeptide or protein fragment.

When transgenic plants and plant seeds are used in accordance with the

30   present invention, they additionally can be treated with the same materials as are used to treat the plants and seeds to which a hypersensitive response eliciting fragment is applied. These other materials, including hypersensitive response eliciting fragments,

- 31 -

can be applied to the transgenic plants and plant seeds by the above-noted procedures, including high or low pressure spraying, injection, coating, and immersion. Similarly, after plants have been propagated from the transgenic plant seeds, the plants may be treated with one or more applications of the hypersensitive response eliciting
5   fragment to impart disease resistance, enhance growth, and/or control insects. Such plants may also be treated with conventional plant treatment agents (e.g., insecticides, fertilizers, etc.).

EXAMPLES

10

Example 1 - Strains and plasmids used

The strains and plasmids used are set forth in Table 1 below

15   Table 1

Plasmid name	E. amylovora source strain	Brief Description, Relevant Phenotype, Reference	Harpin, fragment (or NA)
pBCKS	-	Cmʳ derivative of pBluescript KS. Stratagene, La Jolla, CA	
pBCSK	-	Cmʳ derivative of pBluescript SK. Stratagene, La Jolla, CA	
pBSKS	-	pBluescript KS; Apʳ. Stratagene, La Jolla, CA	
pBSSK II	-	pBluescript II SK; Apʳ. Stratagene, La Jolla, CA	
pBW7	-	Mob+Tcʳ helper plasmid. (Rella, et al., "Transposon Insertion Mutagenesis of *Pseudomonas aeruginosa* With a Tn5 Derivative: Application to Physical Mapping of the *arc* Gene Cluster," Gene 33:293-303 (1985), which is hereby incorporated by reference)	NA
pCPP51	-	Broad host range derivative of pBSSK II containing *ori* from pRO1614.	NA
PCPP430	321	Functional *hrp* gene cluster of *E. amylovora* Ea321 cloned into pCPP9. Beer, S.V., et al., "The *hrp* Gene Cluster of *Erwinia Amylovora*," *In* Hennecke, H., and D.P.S. Verma, (eds.), Advances in Molecular Genetics of Plant-Microbe Interactions, Kluwer Academic Publishers, Dordrecht, Netherlands, 53-60 (1991), which is hereby incorporated by reference.	
pCPP460	246	Functional *hrp* gene cluster of *E. amylovora* Ea246 cloned into pCPP9.	
pCPP1104	321	1.2 kb *Pstl* fragment of pCPP1084 in pBCKS	Ea C31
pCPP1105	321	1.1 kb *Smal* fragment of pCPP1084 in pBCSK	Ea C82
pCPp1107	321	1.3 kb *Hindlll* fragment of pCPP1084 in pBCSK	Ea wt
pCPp1108	321	1.2 kb *Hincll-Hindlll* of pCPP1084 in pBCSK	Ea N11
pCPP1109	321	pCPP1107 with internal *Avall* fragment deleted	Ea 1175
pCPP1110	321	As pCPP1108, but cloned into pBCKS	Ea N9
pCPP1111	321	367 bp *Taql* fragment of pCPP1107 in pBCSK	Ea C305

pCPP1113	246	As pCPP1109, but 425bp *AvaII* fragment of pCPP1098 deleted	Ear 1175
pCPP1119	246	Site specific mutation in pCPP1098; stop codon inserted at L36	Ear C368
pCPP1120	246	Site specific mutation in pCPP1098; stop codon inserted at T123	Ear C281
pCPP1121	321	702bp *KpnI* fragment internal to *hrpN* deleted	Ea C375
pCPP1127	246	3.1 kb *BamHI* fragment of pCPP1098 in pSU21	Ear wt
pCPP1128	246	Tn*10* minikan in pCPP1127	Ear undef
pCPP1136	246	4.4 kb *EcoRI* fragment of pCPP1120, religated	EAR N122
pCPP1146	246	4.2 kb *EcoRI* fragment of pCPP1119, religated	Ear N35
pCPP1147	321	1.2 kb *BamHI* fragment of pCPP1084, PCR amplified, cloned into pSU23	Ea wt
pCPP1148	246	As pCPp1147, but from pCPP1098	Ear wt
pCPP1150	246	As pCPp1148, but in pCPP51 vector	Ear wt
pCPP1163	246	3.1 kb *BamHI* fragment of pCPP1098 in pCPP51	Ear wt
pCPP1164	321	1.3 kb *HindIII* of pCPp1084 in pCPP51	Ea wt
pCPP1165		Derivative of pCPP51 w / *KpnI*, *SacII* sites removed	NA
pCPP1167	321	1.3 kb *HindII* fragment of pCPp1107 in pCPP1165	Ea wt
pCPP1169	246	As pCPP1167, but 3.1 kB *BamHI* insert from pCPp1098	Ear wt
pCPP1170	246	PCPP1098; Σ-Sp ligated into *EcoRV* site	Ear C128Σ
pCPP1171	246	*KpnI* fragment internal to *hrpN* deleted; shifted reading frame	Ea C375
pCPP1172	321	*Derivative of pCPP1167 with in-frame deletion of* KpnI *fragment internal to* hrpN	Ea 1235
pCPP1173	246	*As pCPP1172, but from pCPP1169*	Ear 1235
pCPP217	321	*PCPP1084 with 2 StyI fragments deleted, blunted, and religated*	Ea C185
pCPP1252	321	*PCPP1105 with Σ-Sp ligated at SmaI site*	Ea C82Σ
pGP1-2		*P15a ori.T7* RNA polymerase-encoding plasmid; for protein expression (Tabor, S., et al., "A Bacteriophage T7 DNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes," Proc. Natl. Acad. Sci. USA, 82:1074-1078 (1985), which is hereby incorporated by reference).	N/A
pHP45Σ Ω		Ap^r; Sp^r; source of Ω-Sp fragment; (Fellay, R., et al., "Interposon Mutagenesis of Soil and Water Bacteria a Family of DNA Fragments Designed for *in vitro* Insertional Mutagenesis of Gram-Negative Bacteria," Gene, 52:147-154 (1987), which is hereby incorporated by reference).	N/A
pSU21		P15a *ori* Km^r(Bartolomé, B.Y., et al., "Construction and properties of a Family of pACYC184-Derived Cloning Vectors Compatible With pBR322 and its Derivatives," Gene, 102:75-78 (1991), which is hereby incorporated by reference).	N/A
PSU23		P15a *ori* Km^r(Bartolomé, B.Y., et al., "Construction and properties of a Family of pACYC184-Derived Cloning Vectors Compatible With pBR322 and its Derivatives," Gene, 102:75-78 (1991), which is hereby incorporated by reference).	N/A
Strains used			
E. amylovora		Ea273Nx; Nalidixic acid resistant (Nx^r) derivative of Ea273. CUCPB 2348	
E. amylovora		Rifampin resistant derivative of Ea32. CUCPB 2545	

E. coli	GM272; *dam-, dcm-.* CUCPB 3047; (Blumenthal, R.M., et al., "*E. coli* Can Restrict Methylated DNA and May Skew Genomic Libraries," Trends in Biotech, 4:302-305 (1986), which is hereby incorporated by reference)
E. coli	BL21(DE3); CUCPB 4277; (Studier, F.W., and B.A. Moffatt, "Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes," J. Mol. Biol., 189:113-130 (1986), which is hereby incorporated by reference)
E. coli	DH5α; (Nx^r). CUCPB 2475; Stratagene, La Jolla, CA.

Example 2 - Molecular biology techniques.

Several approaches were employed to obtain truncated or otherwise altered versions of both *E. amylovora* harpins. These techniques included: (i) subcloning of restriction fragments containing portions of the gene encoding the hypersensitive response elicitor protein or polypeptide from *Erwinia amylovora* (i.e. *hrpN*) into expression vectors, by standard techniques (Sambrook, et al., Molecular Cloning: a Laboratory Manual, 2nd ed. ed. Cold Spring Harbor, Laboratory," Cold Spring Harbor, NY (1989), which is hereby incorporated by reference); (ii) insertion of an Ω-fragment (Fellay, et al., "Interposon Mutagenesis of Soil and Water Bacteria a Family of DNA Fragments Designed for *in vitro* Insertional Mutagenesis of Gram-Negative Bacteria," Gene 52:147-154 (1987), which is hereby incorporated by reference) into *hrpN*; (iii) site-specific mutagenesis approaches (Innis, et al., PCR Protocols. A Guide to Methods and Applications, Academic Press San Diego, CA (1990); Kunkel, et al., "Rapid and Efficient Site-Specific Mutagenesis Without Phenotypic Selection," Proc. Nat. Acad. Sci. USA 82:488-492 (1985), which are hereby incorporated by reference); and (iv) creation of nested deletions (Erase-a-Base™ kit; Promega, Madison, WI). C-terminal deletion analysis of the hypersensitive response elicitor protein or polypeptide from *Erwinia amylovora* (i.e. harpin$_{Ea}$) in pCPP1084 could not be performed because of the location of restriction enzyme cleavage sites in pCPP1084. For N-terminal deletions, pCPP1084 DNA was prepared using a Qiagen midiprep column (Qiagen, Chatsworth, CA) and digested with *sst I* followed by *EcoRI*. Subsequently, the digested DNA was subjected to exonuclease III digestion, ligation, and transformation into *E. coli* BL21(DE3). Deletion sizes were estimated by agarose gel electrophoresis. Harpin fragments were named with respect to the portion of harpin deleted (e.g., harpin$_{Ea}$ C82 lacks the C-terminal 82 amino acid residues of full-length harpin$_{Ea}$).

- 34 -

Example 3 - Protein expression.

For expression from T7 promoters, T7 RNA polymerase-dependent systems were used. These systems utilized either strain *E. coli* BL21(DE3) (Studier, et al.,

5   "Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Gene," J. Mol. Biol. 189:113-130 (1986), which is hereby incorporated by reference), or plasmid pGP1-2 (Tabor, et al., "A Bacteriophage T7 DNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes," Proc. Natl. Acad. Sci., USA 82:1074-1078 (1985), which is hereby

10   incorporated by reference) in *E. coli* DH5α. Expression of *hrpN* from the T7 promoter was induced by addition of IPTG to a final concentration of 0.4 mM. For expression in *E. amylovora* Ea321 (i.e. harpin$_{Ea}$) or Ea273, pGP1-2 was introduced by transformation with a 42°C heat shock for 10 minutes, or by electroporation (Biorad Gene Pulser^{TM}). Hypersensitive response (i.e. HR)-eliciting activity was screened in

15   tobacco cv. Xanthi leaves by *in planta* lysis (He, et al., "*Pseudomonas syringae* pv. *syringae* harpin$_{Pss}$: a Protein That is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference) or by preparation of boiled and unboiled "CFEPs" (Wei, et al., "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen

20   *Erwinia amylovora*," Science 257:85-88 (1992), which is hereby incorporated by reference).

Example 4 - *In vitro* proteolysis of harpin.

25   *In vitro* proteolysis of harpin$_{Ea}$ with *Staphylococcus* V8 proteinase (also termed endoproteinase Glu-C), trypsin, pepsin, and papain was performed as recommended (Scopes, et al., Protein Purification: Principles and Practice, 2nd ed. Springer-Verlag. New York (1987), which is hereby incorporated by reference), for 2-16 hrs. at 20-37°. Endoproteinase Glu-C digestion was performed either in 50 mM

30   ammonium bicarbonate, pH 7.8 (in which cleavage occurs only after glutamic acid), or in 50 mM potassium phosphate, pH 7.8 (in which cleavage after both glutamic acid and aspartic acid occurs).

- 35 -

Example 5 - Plant-derived proteinases.

Intercellular fluids (IF) were obtained from tobacco, tomato, apple, raspberry, and cotoneaster, as described (Hammond-Kosack, et al., "Preparation and Analysis of

5   Intercellular Fluid," p. 15-21. *In* S.J. Gurr, M.J. McPherson, and D.J. Bowles (ed.), Molecular Plant Pathology A Practical Approach, 2nd ed., The Practical Approach Series, IRL Publishers, Oxford (1992), which is hereby incorporated by reference), by vacuum infiltration of intercellular spaces with high-purity water. Proteolytic digestion of PAGE-purified harpin$_{Ea}$ was performed for 2-16 hrs. at 20-37°C, pH, by

10   mixing equal volumes of IF with harpin$_{Ea}$. A total leaf extract was obtained by grinding tobacco leaf panels with mortar and pestle in 5 mM potassium phosphate. The extract was centrifuged and filtered, and the clarified ground leaf extract used identically as was the IF. Proteinase inhibitors were employed as follows: Pepstatin A (final concentration 1μM), E-64 (1μM), Aprotinin (2μg/ml), o-phenanthroline

15   (1mM), and p-mercuribenzoate (PCMB) (Sigma, St. Louis, MO).

Example 6 - Peptide purification.

Peptide fragments of harpin obtained following digestion with tobacco IF were

20   fractionated by reverse-phase HPLC on a Vydac C18 column using a 2-60% acetonitrile gradient in 0.1% trifluoroacetic acid. Fractions were lyophilized, resuspended in 5 mM potassium phosphate and infiltrated into tobacco leaf panels. The fraction with greatest HR-eliciting activity was refractionated as above with a 35-70% acetonitrile gradient, and the purity of each fraction was assayed via gas

25   chromatography-mass spectroscopy (GC-MS) and by N-terminal protein sequencing at the Cornell Biotechnology Program Core Facility.

Example 7 - Proteinase activity-stained gels.

30   Proteinase activity of IF was assayed in activity-stained polyacrylamide gels (Laemmli, "Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4," Nature 227:680-685 (1970), which is hereby incorporated by reference) copolymerized with 0.1% gelatin (Heussen, et al., "Electrophoretic Analysis of Plasminogen Activators in Polyacrylamide Gels Containing Sodium

- 36 -

Dodecyl Sulfate and Copolymerized Substrates," Anal. Biochem. 102:196-202 (1980), which is hereby incorporated by reference). After electrophoresis, each gel was rinsed extensively to remove SDS and allow refolding of proteinases in the gel. Following additional incubation to allow proteolysis to occur, the gels were stained

5   with 0.1% Amido Black in 30% methanol/10% acetic acid. Each gel stained darkly (due to the presence of copolymerized gelatin) except where proteinases had digested the gelatin, resulting in colorless bands representing the sites of proteinase activity.

Example 8 - Truncated harpins retain HR-eliciting activity.

10   The stability and the HR-eliciting activity of proteins encoded by various DNA constructs is shown in Figure 1. Many DNA constructs encoding portions of harpin$_{Ea}$ or harpin$_{Ear}$ did not yield detectable protein products following induction of expression in the T7 promoter-polymerase system (Tabor, et al., "A Bacteriophage T7

15   DNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes," Proc. Natl. Acad. Sci. USA 82:1074-1078 (1985), which is hereby incorporated by reference) and analysis of cell extracts by PAGE, possibly due to instability of the encoded proteins. No DNA constructs (e.g., those obtained via Erase-a-Base™ protocol) yielded detectable protein products displaying N-terminal

20   deletions relative to the full-length protein. No stable but inactive proteins were identified. Several constructs encoding proteins truncated at their C-terminus and often including additional vector-encoded amino acids yielded detectable products (e.g. harpin$_{Ea}$ C82). In contrast, a construct encoding the same 321 N-terminal amino acid residues of harpin$_{Ea}$, but yielding a protein truncated by the presence of an Ω-

25   fragment (harpin$_{Ea}$ C82 Ω) was unstable (i.e. no product was detected). A construct encoding a harpin$_{Ea}$ fragment with a large internal deletion (harpin$_{Ea}$ I175) was also successfully used to express protein. These various truncated proteins were tested for HR-eliciting activity. A 98 residue N-terminal harpin$_{Ea}$ fragment (harpin$_{Ea}$ C305) was the smallest bacterially-produced peptide that displayed HR-eliciting activity.

30

- 37 -

Example 9 - Secretion of harpin_Ea with an altered C-terminus.

The effect of alteration at the harpin C-terminus on its secretion was examined. Harpin C31 contains the N-terminal 372 amino acids of harpin, but lacks the C-terminal 31 residues, which are replaced by 47 residues encoded by the vector, resulting in a protein slightly larger than the wild type harpin_Ea. The C31 protein retains HR-eliciting activity and is stable and easily expressed and detected by western analysis or PAGE but it is no longer secreted into the culture supernatant as is the wild type protein (Figure 2). The presence of harpin_Ea C31 does not interfere with secretion of the wild type harpin, which is found in both the CFEP and the culture supernatant. However, harpin_Ea C31 is found only in the CFEP.

Example 10 - Effect of proteolysis on harpin_Ea's HR eliciting activity

In order to generate additional harpin_Ea fragments, purified full length protein was proteolyzed in vitro by several proteinases, including endoproteinase Glu-C, trypsin, pepsin, and papain (e.g., Figures 3 and 4). Harpin solutions digested with trypsin or with papain lost all activity. In contrast, following digestion with endoproteinase Glu-C, HR-eliciting activity was retained. No peptides larger than 6 kD were evident by PAGE following trypsin digestion. Endoproteinase Glu-C digestion yielded an approximately 20 kD fragment, larger than expected if all cleavage sites were cut, indicating that digestion was not complete (Figure 4).

Example 11 - Apoplastic fluids (IF) contain harpin-degrading proteolytic activity

Apoplastic fluids (intercellular fluids; IF) from tobacco and other plants were also employed to proteolyze harpin. Each IF tested possessed proteinase activity(s), as indicated by the presence of multiple activity-stained bands in polyacrylamide gels containing co-polymerized gelatin (Figures 5A to 5C), as well as by the disappearance of detectable harpin_Ea (Schägger, et al., "Tricine-Sodium Dodecyl Sulfate Gel Electrophoresis for the Separation of Proteins in the Range From 1 to 100 kDa," Anal. Biochem. 166:368-379 (1987), which is hereby incorporated by reference) following overnight digestion of purified harpin_Ea with IF. Proteinase activity was substantially greater at 37°C than at 20°C, and activity was higher at pH 8.5 than at pH 7. Several

inhibitors were employed in order to define the proteolytic activity(s) of the IF. No single proteinase inhibitor which was employed prevented degradation of harpin$_{Ea}$. However, a mixture of the inhibitors Pepstatin A (1μM), E-64 (1μM), Aprotinin (2μg/ml), and o-phenanthroline (1mM), targeted at acid proteinases, cysteine

5   proteinases, serine proteinases, and metalloproteinases, respectively, partially inhibited proteolysis.

Harpin$_{Ea}$ degraded by proteolytic activities present in the plant apoplast retained HR-eliciting activity (Figure 3). In contrast, harpin$_{Ea}$ proteolyzed by a clarified extract produced by grinding tobacco leaf tissue with mortar and pestle lost

10   HR-eliciting activity. In order to study whether apoplastic degradation of harpin was a prerequisite to its HR-eliciting activity, the length of time required for leaf collapse when either intact harpin or harpin predigested with tobacco IF was infiltrated into tobacco leaf panels was compared. Both preparations elicited the HR in a similar time frame (12-18 hours, depending on the experiment).

15

Example 12 - Characterization of HR-eliciting peptide fragments

Peptides resulting from digestion by apoplastic plant proteinase(s) were fractionated by reverse phase HPLC (Vydac C18 column), and tested for activity.

20   Following treatment of intact harpin$_{Ea}$ with tobacco IF, three fractions contained some HR-eliciting activity on tobacco. Two of the three demonstrated weak activity, and little protein was present. They were not further characterized. Fraction 19, which contained the strongest activity as well as the most protein, was refractionated using a more shallow elution gradient (Figure 6). Refractionation, N-terminal protein

25   sequencing, and molecular weight analysis by mass spectroscopy indicated that four largely overlapping peptides were present. Peak 19-1 contained peptides P91 and P95, corresponding to harpin$_{Ea}$ residues 110-200 and 110-204; peak 19-2 contained peptides P64 and P68, corresponding to harpin$_{Ea}$ residues 137-200 and 137-204. 19-1 and 19-2 each possessed HR-eliciting activity. The smallest peptide thus confirmed

30   to retain activity consisted of residues 137-204. The two peptides in each peak were not separable under the conditions used. These active fragments are distinct from the smallest active N-terminal fragment (harpin$_{Ea}$C305), and indicate that more than one portion of harpin$_{Ea}$ displays activity *in planta*. Further digestion with trypsin

abolished the HR-eliciting activity of 19-2. This proteinase cleaves P64 and P68 as shown in Figure 7. Further digestion with endoproteinase Glu-C in ammonium bicarbonate buffer abolished the HR-eliciting activity of 19-1. Endoproteinase Glu-C is predicted to cleave P91 and P95 as shown in Figure 7. Loss of elicitor-activity

5 followed further digestion of these peptides with endoproteinase Glu-C or trypsin.

Example 13 - *E. amylovora* harpin's similarity with other proteins.

The predicted protein sequences of proteinaceous HR elicitors from several
10 other bacterial plant pathogens, and of other proteins known to be, or thought to be, secreted by a type III secretion pathway were also compared with that of harpin$_{Ea}$. When harpin$_{Ea}$ was compared with elicitors from *E. amylovora* Ea246 (i.e. harpin$_{Ear}$), *Erwinia chrysanthemi* EC16 (harpin$_{Ech}$) (Bauer, et al., "*Erwinia chrysanthemi* harpin$_{Ech}$: An Elicitor of the Hypersensitive Response That Contributes to Soft-Rot
15 Pathogenesis," Mol. Plant-Microbe Interact 8:484-491 (1995), which is hereby incorporated by reference), *Erwinia carotovora* subsp. *carotovora* (harpin$_{Ecc}$) (Mukherjee, et al., Presented at the 8th International Congress Molecular Plant-Microbe Interactions, Knoxville, TN (1996), which is hereby incorporated by reference), *Erwinia stewartii* (Harpin$_{Es}$) (Frederick, et al., "The *wts* Water-Soaking
20 Genes of *Erwinia stewartii* are Related to *hrp* genes," Presented at the Seventh International Symposium on Molecular Plant-Microbe Interactions, Edinburgh, Scotland (1994), which is hereby incorporated by reference), *Ralstonia (Pseudomonas) solanacearum* (PopA) (Arlat, et al., "PopA1, a Protein Which Induces a Hypersensitivity-Like Response on Specific *Petunia* Genotypes, is Secreted via the
25 Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-553 (1994), which is hereby incorporated by reference), *Pseudomonas syringae* 61 (harpin$_{Pss}$) (He, et al., "*Pseudomonas syringae* pv. *syringae* harpin$_{Pss}$: a Protein That is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference), *Pseudomonas syringae* pv.
30 *tomato* (harpin$_{Pst}$) (Preston, et al., "The HrpZ Proteins of *Pseudomonas syringae* pvs. *syringae, glycinea,* and *tomato* Are Encoded By An Operon Containing *Yersinia ysc* Homologs and Elicit the Hypersensitive Response in Tomato But Not Soybean," Mol. Plant-Microbe Interact 8:717-732 (1995), which is hereby incorporated by reference),

the *Erwinia*-derived harpins contained significant regions of similarity at the C-terminus. In addition, all the elicitors are glycine-rich, secreted, and heat-stable. Limited similarity between harpin$_{Pss}$ and harpin$_{Ea}$ had been reported previously (He, et al., "*Pseudomonas syringae* pv. *syringae* harpin$_{Pss}$: a Protein That is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference), (Laby, et al., Presented at the Seventh International Symposium on Molecular Plant-Microbe Interactions, Edinburgh, Scotland (1994), which is hereby incorporated by reference). A limited region of similarity between harpin$_{Ea}$ and other harpins frm *Erwinia* spp. was also evident at the extreme N-terminus of each protein, where 9 out of the first 26 residues are conserved (Figure 8). Kyte-Doolittle hydropathy plots for each of the harpins produced by different *Erwinia* spp. are illustrated in Figure 9. Each of the *Erwinia* harpins examined displays a generally similar hydrophobicity profile along the full length of the protein. This profile is distinct from the profile demonstrated by PopA1 and by harpin$_{Pss}$, and does not possess the symmetry evident in the profile of those two proteins. The hydropathy profile of each *Erwinia* harpin is generally similar to that of the others, yet distinct from that reported for harpin$_{Pss}$ (Alfano, et al., "Analysis of the Role of the *Pseudomonas Syringae* HrpZ harpin in Elicitation of the Hypersensitive Response to Tobacco Using Functionally Nonpolar *hrpZ* Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations," Mol. Microbiol. 19:715-728 (1996), which is hereby incorporated by reference). Harpin$_{Ecc}$ possesses a strikingly hydrophobic region around residues 54-143 (Mukherjee, et al., Presented at the 8th International Congress Molecular Plant-Microbe Interactions, Knoxville, TN (1996), which is hereby incorporated by reference). This portion of the protein is also the most hydrophobic region of harpin$_{Ea}$ and harpin$_{Es}$. The rest of each protein is predominantly hydrophilic.

Truncated proteins and fragments of harpin obtained following proteolytic digestion of the full length protein indicate several surprising aspects of harpin$_{Ea}$ HR-eliciting activity. These harpin fragments demonstrate that HR-eliciting activity resides in distinct regions of the protein, and that relatively small fragments of the protein, as little as 68 residues and possibly less, are sufficient for this activity. Fragments of other plant pathogen-derived elicitor proteins also retain biological

- 41 -

activity, including Avr9 from *Caldosporium fulvum* (Van den Ackervecken, et al.,
"The AVR9 Race-Specific Elicitor of *Cladosporium fulvum* is Processed by
Endogenous and Plant Proteases," Pl. Physiol. 103:91-96 (1993), which is hereby
incorporated by reference), Pep-13 of *Phytophthora megasperma* (Nürnburger, et al.,

5   "High Affinity Binding of a Fungal Oligopeptide Elicitor to the Parsley Plasma
Membranes Triggers Multiple Defence Responses," Cell, 78:449-460 (1994), which is
hereby incorporated by reference), and harpin$_{Pss}$ of *P. syringae* pv. *syringae* (Alfano,
et al., "Analysis of the Role of the *Pseudomonas syringae* HrpZ harpin in Elicitation
of the Hypersensitive Response in Tobacco Using Functionally Nonpolar *hrpZ*

10   Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations," Mol.
Microbiol. 19:715-728 (1996), which is hereby incorporated by reference).

        Expression of truncated harpin fragments and proteolysis of full-length
harpins showed that two distinct fragments retain HR-eliciting activity. The primary
sequence of each active fragment show no discernable similarity with each other, or

15   with other elicitor-active peptides. However, the sites of cleavage by trypsin and
endoproteinase Glu-C suggest portions of each fragment required for activity. It
would be of interest to alter specifically the amino acid residues at or near these
cleavage sites to determine whether HR-eliciting activity is altered or lost.
Additionally, harpin$_{Ea}$ P64 and P68 demonstrate distinct hydrophobicity during

20   reverse-phase HPLC (Figure 6), and they correspond to a hydrophobic peak in a Kyte-
Doolittle plot (Figure 9). The role of this putative hydrophobic domain could be
tested by mutagenesis, or by synthesis of altered peptides. However, the fact that
multiple fragments independently possess HR-eliciting activity complicates analysis
of full-length proteins.

25   This finding, that fragments of the protein retain HR-eliciting activity, also
allowed (at least) two apoplastic proteinase activities, which are distinct from
intracellular plant proteinases, to be defined. Two apoplastic plant proteinases (from
soybean) have been studied in some detail. SMEP, a metalloproteinase (Huangpu, et
al., "Purification and Developmental Analysis of an Extracellular Proteinase From

30   Young Leaves of Soybean," Plant Physiol 108:969-974 (1995); McGeehan, et al.,
"Sequencing and Characterization of the Soybean Leaf Metalloproteinase," Plant
Physiol. 99:1179-1183 (1992), which are hereby incorporated by reference) sensitive

to EDTA, is thought to cleave at G/L and G/I. Interestingly, although there are 19 potential SMEP cleavage sites in the intact harpin$_{Ea}$, only one of them is located within fragments P91 and P95, and none are within fragments P64 and P68 (Figure 7). P91 and P95 thus may represent partial digestion products of a SMEP-like proteinase

5   in the tobacco apoplast. The other studies soybean apoplastic proteinase, SLAP, a sulfhydryl proteinase (Huangpu, et al., "Purification and Developmental Analysis of an Extracellular Proteinase From Young Leaves of Soybean," Plant Physiol. 108:969-974 (1995), which is hereby incorporated by reference) sensitive to p-chloromercuribenzoic acid (pCMB). Several lines of evidence suggest that multiple

10   proteolytic activities in the IF are degrading harpin$_{Ea}$. PMSF, a serine protease inhibitor, decreases but does not entirely block harpin$_{Ea}$ degradation (Figure 5C); no single proteinase inhibitor tested blocks harpin degradation, and the cleavage sites after residues 109, 136, 200, and 204 are dissimilar. Endoproteinase Glu-C does not abolish activity of full-length harpin, but does abolish activity of P91 and P95 (and

15   presumably P64 and P68); trypsin abolishes the activity of P64 and P68 (and presumably P91 and P95). These final digests suggest specific portions of each distinct HR-eliciting peptide which are presumably necessary for its activity, as mentioned previously.

The apoplastic activities degrade harpin without destroying its HR-eliciting

20   ability, in contrast to intracellular proteolytic activities present in a ground leaf-extract, which abolish activity. This raises a number of intriguing questions, e.g., does the plant use these harpin fragments as elicitor-signals? The timing of the HR was examined when full length harpin and harpin predigested by tobacco intercellular fluid were each infiltrated into tobacco leaves. The HR elicited by each preparation

25   occurred 12-18 hours after infiltration. Co-infiltration of proteinase inhibitors into tobacco leaf panels along with harpin also had no effect on harpin's HR eliciting activity, although limited proteolytic degradation cannot be ruled out in this case, particularly since it seems that at least two, and perhaps more, apoplastic proteinases are present in tobacco. Because predigested harpin elicited the HR no faster than

30   undigested protein, proteolytic digestion seems unlikely to be a rate-limiting step required for HR to occur. The role of these apoplastic proteinases which are able to hydrolyze harpin partially, yet unable to abolish harpin's HR-eliciting activity on

tobacco, remains unclear. Salzer et al., "Rapid Reactions of Spruce Cells to Elicitors Released From the Ectomycorrhizal Fungus *Hebeloma crustuliniforme* and Inactivation of These Elicitors by Extracellular Spruce Cell Enzymes," Planta 198:118-126 (1996), which is hereby incorporated by reference, have noted that

5   spruce (*Picea abies* (L.) Karst.) modulates the level of fungal cell wall elicitors released by the ectomycorrhizal fungus *Hebeloma crustuliniforme* by inactivating these molecules in the apoplast. They propose that *Picea* controls the elicitor level as part of its symbiotic interaction with the fungus. Similarly, PGIP of *Phaseolus vulgaris* has been suggested to modulate the level of elicitor-active

10 oligogalacturonides present during the plant-parasite interaction in bean (Desiderio, et al., "Polygalacturonase, PGIP, and Oligogalacturonides in Cell-Cell Communication," Biochem. Sci. Trans. 22:394-397 (1994), which is hereby incorporated by reference). Perhaps the retention of HR-eliciting activity by harpin fragments contributes to the ability of plants to recognize the presence of a pathogen. In this regard, it would be

15 interesting to explore whether transgenic host and non-host plants, engineered for apoplastic expression of a harpin activity-degrading proteinase, would exhibit reduced or increased sensitivity to *E. amylovora*, compared to non-engineered plants.

Despite numerous attempts, only a handful of truncated derivatives of harpin$_{Ea}$ and harpin$_{Ear}$ were successfully expressed from portions of *hrp*N. Problems with

20 protein stability were evident in that several truncated harpins were unstable and difficult to purify. Additionally, expression of truncated harpins may be deleterious to bacteria. Truncated harpin$_{Ea}$C31 was, however, stable and easily purified, but not secreted, suggesting that C-terminal sequences are involved in harpin secretion. Unfortunately, the presence of vector-encoded amino acids in this protein complicates

25 this conclusion. All attempts to clone β-galactosidase-harpin fusion proteins have been unsuccessful, as were attempts to clone and express *hrp*N downstream of the *lacZ* promoter in several commonly used vectors such as pBluescript. Expression of such constructs is evidently deleterious to bacterial strains.

Wei, et al., "Harpin, Elicitor of the Hypersensitive Response Produced By the

30 Plant Pathogen *Erwinia amylovora*," Science, 257:85-88 (1992), which is hereby incorporated by reference, reported previously that BLAST searches indicated harpin$_{Ea}$ possessed slight similarity to several other glycine-rich proteins, including

keratins and glycine-rich cell wall proteins. However, this is thought to be due to the high glycine content of harpin$_{Ea}$, and as such does not suggest a role for harpin$_{Ea}$. Examination of N-terminal sequences from several HR-eliciting proteins produced by phytopathogenic bacteria (Figure 8) yielded some potential similarities. However, the region in question is quite short. The region of putative primary sequence similarity is limited to the first 26 residues at the N-terminus, and its role remains unclear. Surprisingly, *E. carotovora* harpin$_{Ecc}$ appears more similar to the harpins from *E. amylovora* and *E. stewartii* than to that from *E. chrysanthemi*, to which it is more closely related with respect to its taxonomic position as well as its mechanism of pathogenicity (i.e. soft-rots). In addition, although primary sequence similarity is strongest only in the C-terminal third of each protein, the *Erwinia* harpins possess broadly similar hydrophobicity profiles along their entire lengths (Figure 9). Based on its hydrophobicity profile, Alfano, et al., "Analysis of the Role of the *Pseudomonas syringae* HrpZ harpin in Elicitation of the Hypersensitive Response in Tobacco Using Functionally Nonpolar *hrpZ* Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations," Mol. Microbiol. 19:715-728 (1996), which is hereby incorporated by reference, speculated that harpin$_{Pss}$ may have an amphiphilic nature. However, the *Erwinia* harpins' profiles do not match that of harpin$_{Pss}$.

Recently, a number of other secreted glycine-rich pathogenicity associated proteins, elicitors of the HR or other plant-defense responses, have been described from other plant pathogenic bacteria and fungi (Boller, "Chemoperception of Microbial Signals in Plant Cells," Ann. Rev. Plant Physiol. Plant Molec. Biol. 46:189-214 (1996), which is hereby incorporated by reference), including *Phytophthora megasperma* (Ballieul, et al., "A New Elicitor of the Hypersensitive Response in Tobacco: a Fungal Glycoprotein Elicits Cell Death, Expression of Defence Genes, Production of Salicylic Acid, and Induction of Systemic Acquired Resistance," Plant Journal 8:551-560 (1995); Nürnburger, et al., "High Affinity Binding of a Fungal Oligopeptide Elicitor to the Parsley Plasma Membranes Triggers Multiple Defence Responses," Cell 78:449-460 (1994), which are hereby incorporated by reference), and *Magnaporthe grisea* (Sweigard, et al., "Identification, Cloning, and Characterization of PWL2, a Gene For Host Species Specificity in the Rice Blast Fungus," Plant Cell 7:1221-1233 (1995), which is hereby incorporated by reference).

- 45 -

Proteinaceous HR-elicitors have also now been described from *Phynchosporium secalis* (Rohe, et al., "The Race-Specific Elicitor, NIP1, From the Barley Pathogen, *Rhynchosporium secalis*, Determines Avirulence on Host Plants of the *Rrs*1 Resistance Genotype," EMBO Journal 14:4168-4177 (1995) which is hereby
5   incorporated by reference, while *P. infestans* (Pieterse, et al., "Structure and Genomic Organization of the *ipi*B and *ipi*O Gene Clusters of *Phytophthora infestans*," Gene, 138:67-77 (1994), which is hereby incorporated by reference) produces a glycine-rich pathogenicity-associated family of proteins of unknown function. Because the primary amino acid sequence of each elicitor protein or peptide fragment shows no
10 obvious similarity to that of the others, it is unclear whether they interact with the same target on or in the plant cell, plasma membrane, or cell wall. In that regard, it might be of interest to test whether any one of these molecules inhibits the action of other(s). The increasing availability of peptides such as Pep13, Avr9, P68, and harpin$_{Ea}$ C305 with plant-defense response-eliciting activity (HR and otherwise)
15 should enable precise probing of their targets on or in plant cells, as well as determination of whether their mechanisms of activity are similar, distinct, or overlapping.

Example 14 - Bacterial strains and plasmids

20
Escherichia coli stains used in the following examples include DH5α and BL21(DE3) purchased from Gibco BRL and Stratagene, respectively. The pET28(b) vector was purchased from Novagen. Eco DH5α/2139 contained the complete hrpN gene. The 2139 construct was produced by D. Bauer at the Cornell University. The
25 hrpN gene was cleaved from the 2139 plasmid by restriction enzyme digestion with HindIII, then purified from an agarose gel to serve as the DNA template for PCR synthesis of truncated hrpN clones. These clones were subsequently inserted into the (His)$_6$ vector pET28(b) which contained a Kan^r gene for selection of transformants.

30 ### Example 15 - DNA Manipulation

Restriction enzymes were obtained from Boehringer Mannheim or Gibco BRL. T4 DNA ligase, Calf Intestinal Alkaline Phosphatase (CIAP), and PCR

- 46 -

Supermix™ were obtained from Gibco BRL. The QIAprep Spin Miniprep Kit, the Qiagen Plasmid Mini Kit, and the QIAquick PCR Purification Kit were purchased from Qiagen. The PCR primers were synthesized by Lofstrand Labs Limited (Gaithersburg, MD). The oligopeptides were synthesized by Bio-Synthesis, Inc.

5   (Lewisville, TX). All DNA manipulations such as plasmid isolation, restriction enzyme digestion, DNA ligation, and PCR were performed according to standard techniques (molecular cloning) or protocols provided by the manufacturer.

Example 16 - Fragmentation of hrpN gene
10

A series of N-terminal and C-terminal truncated *hrpN* genes and internal fragments were generated via PCR (Fig. 10). The full length hrpN gene was used as the DNA template and 3' and 5' primers were designed for each truncated clone (Fig. 11). The 3' primers contained in the NdeI enzyme cutting site which contained

15   the start codon ATG (Methionine) and the 5' primers contained the stop codon TAA and a HindIII enzyme cutting site for ligation into the pET28(b) vector. PCR was carried out in 0.5 ml tubes in a GeneAmp™ 9600 or 9700. 45 μl of Supermix™ were mixed with 20 pmoles of each pair of DNA primers, 10 ng of full length harpin DNA, and diH$_2$O to a final volume of 50 μl. After heating the mixture at 95°C for 2 min, the

20   PCR was performed for 30 cycles at 94°C for 1 min, 58°C for 1 min and 72°C for 1.5 min. The PCR products were verified on a 6% TBE gel (Novex). Amplified DNA was purified with the QIAquick PCR purification kit, digested with Nde I and Hind III at 37°C for 5 hours, extracted once with phenol:chloroform:isoamylalcohol (25:25:1) and precipitated with ethanol. 5 μg of pET28(b) vector DNA were digested

25   with 15 units of Nde I and 20 units of Hind III at 37°C for 3 hours followed with CIAP treatment to reduce the background resulting from incomplete single enzyme digestion. Digested vector DNA was purified with the QIAquick PCR purification kit and directly used for ligation. Ligation was carried out at 14-16°C for 5-12 hours in a 15 μl mixture containing ca. 200 ng of digested pET28(b), 30 ng of targeted PCR

30   fragment, and 1 unit T4 DNA ligase. 5 - 7.5 μl of ligation solution were added to 100 μl of DH5α competent cells in a 15 ml falcon tube and incubated on ice for 30 min. After a heat shock at 42°C for 45 seconds, 0.9 ml SOC solution or 0.45 ml LB media

were added to each tube and incubated at 37°C for 1 hour. 20, 100, and 200 μl of transformed cells were placed onto LB agar with 30 μg/ml of kanamycin and incubated at 37°C overnight. Single colonies were transferred to 3 ml LB-media and incubated overnight at 37°C. Plasmid DNA was prepared from 2 ml of culture with

5 the QIAprep Miniprep kit. The DNA from the transformed cells was analyzed by restriction enzyme digestion or partial sequencing to verify the success of the transformations. Plasmids with the desired DNA sequence were transferred into the BL21 strain using the standard chemical transformation method as indicated above. A clone containing the full length harpin protein in the pet28(b) vector was generated

10 as a positive control, and a clone with only the pET28(b) vector was generated as a negative control.

Example 17 - Expression of harpin truncated proteins

15   *Escherichia coli* BL21(DE3) strains containing the hrpN clones were grown in Luria broth medium (g/L Difco Yeast extract, 10 g/L Difco Tryptone, 5g/L NaCl, and 1 mM NaOH) containing 30 μg/ml of kanamycin at 37°C overnight. The bacteria were then inoculated into 100 volumes of the same medium and grown at 37°C to an OD_{620} of 0.6-0.8. The bacteria were then inoculated into 250 volumes of the same

20 medium and grown at 37°C to an OD_{620} of ca. 0.3 or 0.6-0.8. One milli molar IPTG was then added and the cultures grown at 19°C overnight (ca. 18 hours). Not all of the clones were successfully expressed using this strategy. Several of the clones had to be grown in Terrific broth (12 g/L Bacto Tryptone, 24 g/L Bacto yeast, 0.4% glycerol, 0.17 M KH_2PO_4, and 0.72 K_2HPO_4), and/or grown at 37°C after IPTG

25 induction, and/or harvested earlier than overnight (Table 2).

- 48 -

Table 2: Expression of harpin truncated proteins

Fragment	amino acids (SEQ. ID. No. 23)	Growth medium	Induction O.D.	Expression temp.	Harvest time
1 (+ control)	1-403	LB	ca. 0.3 or 0.6-0.8	19°C or 25°C	16-18 hr
2 (+ control)	-	LB and TB	ca. 0.3 or 0.6-0.8	19 C and 37 C	16-18 hr
3	105-403	LB	0.6-0.8	19°C	16-18 hr
4	169-403	TB	ca. 0.3	19°C	16-18 hr
5	210-403	LB or M9ZB	0.6-0.8	19°C	16-18 hr
6	257-403	LB or M9ZB	0.6-0.8	19°C	16-18 hr
7	343-403	LB	ca. 0.3	19°C	5 hr
8	1-75	TB	ca. 0.3	37°C	16-18 hr
9	1-104	TB	ca. 0.3	37°C	16-18 hr
10	1-168	TB	ca. 0.3	37°C	16-18 hr
11	1-266	LB	ca. 0.3	37°C	4 hr
12	1-342	LB	0.6-0.8	19°C	16-18 hr
13	76-209	LB	ca. 0.3	37°C	5 hr
14	76-168	TB or LB	ca. 0.3	37°C	3 hr or 16-18 hr
15	105-209	M9ZB	ca. 0.3	37°C	3 hr
16	169-209	no expression			
17	105-168	LB	ca. 0.3	37°C	3-5 hr
18	99-209	LB	ca. 0.3	37°C	3 hr
19	137-204	LB	ca. 0.3	37°C	3 hr
20	137-180	LB	ca. 0.3	37°C	16-18 hr.
21	105-180	LB	ca. 0.3	37°C	3 hr
22	150-209	no expression			
23	150-180	no expression			

General expression method: Escherichia coli BL21(DE3) strains containing the hrpN subclones were grown in Luria broth medium (5g/L Difco Yeast extract, 10 g/L Difco Tryptone, 5g/L NaCl, and 1 mM NaOH) containing 30 µg/ml of kanamycin at 37°C overnight. The bacteria were then inoculated into 100 volumes of the same medium and grown at 37°C to an OD_{620} of 0.6-0.8. The bacteria were then inoculated into 250 volumes of the growth medium and grown at 37°C to a specific induction OD_{620}. One milli molar IPTG was then added and the cultures grown at an optimal temperature for protein expression, and harvested at a particular time for recovery of the highest level of protein.

Example 18 - Small scale purification of harpin truncated proteins (verification of expression)

A 50 ml culture of a hrpN clone was grown as above to induce expression of the truncated protein. Upon harvesting of the culture, 1.5 ml of the cell suspension were centrifuged at 14,000 rpm for 5 minutes, re-suspended in urea lysis buffer (8 M urea, 0.1 M Na_2HPO_4, and 0.01 M Tris -- pH 8.0), incubated at room temperature for 10 minutes, then centrifuged again at 14,000 rpm for 10 minutes, and the supernatant

- 49 -

saved. A 50 μl aliquot of a 50% slurry of an equilibrated (His)$_6$-binding nickel agarose resin was added to the supernatant and mixed at 4°C for one hour. The nickel agarose was then washed three times with urea washing buffer (8 M urea, 0.1 M Na$_2$HPO$_4$, and 0.01 M Tris -- pH 6.3), centrifuging at 5,000 rpm for five minutes

5   between washings. The protein was eluted from the resin with 50 μl of urea elution buffer (8 M urea, 0.1 M Na$_2$HPO$_4$, 0.01 M Tris, and 0.1 M EDTA -- pH 6.3). The eluate was run on a 4-20%, a 16%, or a 10-20% Tris-Glycine pre-cast gel depending upon the size of the truncated protein to verify the expression.

Example 19 - Induction of HR in tobacco

10   A 1.5 ml aliquot from the 50 ml cultures grown for small scale purification of the truncated proteins was centrifuged at 14,000 rpm for four minutes and re-suspended in an equal volume of 5 mM potassium phosphate buffer, pH 6.8. The cell suspension was sonicated for ca. 30 seconds then diluted 1:2 and 1:10 with

15   phosphate buffer. Both dilutions plus the neat cell lysate were infiltrated into the fourth to ninth leaves of 10-15 leaf tobacco plants by making a hole in single leaf panes and infiltrating the bacterial lysate into the intercellular leaf space using a syringe without a needle. The HR response was recorded 24-48 hr post infiltration. Tobacco (*Nicotiana tabacum* v. Xanthi) seedlings were grown in an environmental

20   chamber at 20-25°C with a photoperiod of 12-h light /12-h dark and ca. 40% RH. Cell lysate was used for the initial HR assays (in order to screen the truncated proteins for HR activity) as the small scale urea purification yielded very little protein which was denatured due to the purification process.

25   ### Example 20 - Large scale native purification of harpin truncated proteins for comprehensive biological activity assays

Six 500 ml cultures of a hrpN clone were grown as described earlier to induce expression of the truncated protein. Upon harvesting of the culture the cells were

30   centrifuged at 7,000 rpm for 5 minutes, re-suspended in imidazole lysis buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris) plus Triton X-100 at 0.05% and lysozyme at 0.1 mg/ml, and incubated at 30°C for 15 minutes, sonicated for two minutes, then centrifuged again at 15,000 rpm for 20 minutes, and the supernatant was saved. A 4

ml aliquot of a 50% slurry of an equilibrated $(His)_6$-binding nickel agarose resin was added to the supernatant and mixed at 4°C for ca. four hours. The nickel agarose was then washed three times with imidazole washing buffer (20 mM imidazole, 0.5 M NaCl, and 20 mM Tris), centrifuging at 5,000 rpm for five minutes between washings,

5   then placed in a disposable chromatography column. The column was centrifuged at 1100 rpm for one minute to remove any residual wash buffer and then the protein was eluted from the resin with 4 ml of imidazole elution buffer (1 M imidazole, 0.5 M NaCl, and 20 mM Tris) by incubating the column with the elution buffer for ten minutes at room temperature and then centrifuging the column at 1100 rpm for one

10   minute. The eluate was run on a 4-20%, a 16%, or a 10-20% Tris-Glycine pre-cast gel depending upon the size of the truncated protein to verify the expression. The concentration of the proteins was determined by comparison of the protein bands with a standard protein in the Mark 12 molecular weight marker.

15   <u>Example 21</u> - **Large scale urea purification of harpin truncated proteins for comprehensive biological activity assay**

The procedure was the same as the large scale native purification except that urea lysis buffer, washing buffer, and elution buffer were used, and the cells were not

20   sonicated as in the native purification. After purification, the protein was renatured by dialyzing against lower and lower concentrations of urea over an eight hour period, then dialyzing overnight against 10 mM Tris/20 mM NaCl. The renaturing process caused the N-terminal proteins to precipitate. The precipitated 1-168 protein was solubilized by the addition of 100 mM Tris-HCl at pH 10.4 then heating the protein at

25   30°C for ca. one hour. The concentration of the protein was determined by comparison of the protein bands with a standard protein in the Mark 12 molecular weight marker. The 1-75 and 1-104 protein fragments were not successfully solubilized using this strategy so they were sonicated in 100 mM Tris-HCl at pH 10.4 to solubilize as much of the protein as possible and expose the active sites of the

30   protein for the biological activity assays.

- 51 -

Example 22 - Expression of harpin truncated proteins

The small scale expression and purification of the fragment proteins was done to screen for expression and HR activity (Table 3).

5

Table 3

Expression and HR activity of harpin truncated proteins (small scale screening)

Fragment #	Amino Acids (SEQ. ID. No. 23)	Expression	HR activity
1 (+control)	1-403	+	+
2 (- control)	-	background protein only	-
3	105-403	+	+
4	169-403	+	-
5	210-403	+	-
6	267-403	+	-
7	343-403	+/-	-
8	1-75	+	-
9	1-104	+	+/-
10	1-168	+	+.
11	1-266	+	+
12	1-342	+	+
13	76-209	+	+
14	76-168	+	-
15	105-209	+	+
16	169-209	-	-
17	105-168	+	-
18	99-209	+	+
19	137-204	+	+
20	137-180	+	+
21	105-180	+	+
22	150-209	-	-
23	150-180	-	-

10   All of the cloned fragment proteins were expressed to a certain degree except for three small fragments (amino acids 169-209, 150-209, and 150-180). The fragments were expressed at varying levels. Fragments 210-403 and 267-403 were expressed very well, yielding a high concentration of protein from a small scale purification, resulting in a substantial protein band on SDS gel electrophoresis. Other fragments (such as

15   a.a. 1-168 and 1-104) produced much less protein, resulting in faint protein bands upon electrophoresis. It was difficult to determine whether fragment 343-403, the smallest C-terminal protein, was expressed, as there were several background proteins apparent on the gel, in addition to the suspected 343-403 protein. The positive and

- 52 -

negative control proteins, consisting of the full length harpin protein and only background proteins, respectively, were tested for expression and HR activity as well.

The large scale expression and purification of the fragment proteins was done to determine the level of expression and titer of the HR activity (Table 4).

5

Table 4

Expression level and HR titer of harpin truncated proteins (large sale purification)

Fragment #	Amino acids (SEQ. ID. No. 23)	Expression	HR titer
1(+ control)	1-403	3.7 mg/ml	5-7 µg/ml
2 (- control)	-	-	1:2 dilution
4	169-403	2 mg/ml	-
5	210-403	5 mg/ml	-
6	267-403	4 mg/ml	-
7	343-402	200µg/ml	-
8	1-75	50µg/ml	-
9	1-104	50µg/ml	3 µg/ml (1:16 dilution)
10	1-168	1 mg/ml	1 µg/ml
13	76-209	2.5 mg/ml	5 µg/ml
14	76-168	2 mg/ml	-
15	105-209	5 mg/ml	5-10µg/ml
17	105-168	250µg/ml	-
19	137-204	3.6 mg/ml	3.5 µg/ml
20	137-180	250 µg/ml	16 µg/ml

10

Not all of the proteins were expressed in large scale due to time constraints. The truncated proteins deemed to be the most important in characterizing harpin were chosen. The positive control (full length harpin) was expressed in a relatively high level at 3.7 mg/ml. All of the C-terminal proteins were expressed at relatively high levels from 2-5 mg/ml, except for fragment 343-403 as discussed earlier. The N-terminal fragments were expressed very well also, however, during the purification process, the protein precipitated and very little was resolubilized. The concentrations in Table 3 reflect only the solubilized protein. The internal fragments were expressed in the range of 2-3.6 mg/ml. It was extremely difficult to determine the concentration of fragment 105-168 (it was suspected that the concentration was much higher than indicated), as the protein bands on the SDS gel were large, but poorly stained. The negative control contained several background proteins as expected, but no obviously induced dominant protein.

- 53 -

Example 23 - Induction of HR in tobacco

The full length positive control protein elicited HR down to only 5-7µg/ml. The negative control (pET 28) imidazole purified "protein" - which contained only background proteins - elicited an HR response down to the 1:2 dilution, which lowered the sensitivity of the assay as the 1:1 and 1:2 dilutions could not be used. This false HR was likely due an affinity of the imidazole used in the purification process to bind to one or several of the background proteins, thereby not completely dialyzing out. Imidazole at a concentration of ca. 60 mM did elicit a false HR response.

One definitive domain encompassed a small internal region of the protein from a.a. 137-180 (SEQ. ID. No. 23), a mere 44 a.a, is identified as the smallest HR domain. The other potential HR domain is thought to be located in the N-terminus of the protein from a.a. 1-104 (possibly a.a. 1-75) (SEQ. ID. No. 23). It was difficult to confirm or narrow down the N-terminus HR domain due to the difficulties encountered in purifying these fragment proteins. The N-terminus fragment proteins had to be purified with urea as no protein was recovered when the native purification process was used. Consequently, these proteins precipitated during the renaturing process and were difficult or nearly impossible to get back into solution, thereby making it hard to run the proteins through the HR assay, as only soluble protein is able to elicit HR. Difficulty narrowing the N-terminus HR domain was only compounded by the fact that the negative control elicited false HR at the low dilution levels thereby reducing the sensitivity of the assay.

The internal domain proteins elicited an HR response between 5 and 10 µg/ml of protein like the positive control, and the N-terminus domain proteins elicited an HR response between 1 and 3 µg/ml, lower than the positive control.

Surprisingly, when the internal HR domain was cleaved between a.a. 168 and 169 (fragments 76-168 and 105-168) (SEQ. ID. No. 23) the fragment lost its HR activity. This suggests that the HR activity of fragment 1-168 (SEQ. ID. No. 23) should not be attributed to the internal HR domain, but rather to some other domain, leading to the assumption that there was likely a second HR domain to be found in the N-terminal region of the protein. However, as discussed earlier it was difficult to confirm this assumption.

- 54 -

The harpin C-terminus (a.a. 210-403 (SEQ. ID. No. 23)) did not contain an HR domain. It did not elicit HR at a detectable level using the current HR assay. Even the large C-terminal fragment from a.a. 169-403 (SEQ. ID. No. 23) did not elicit HR even though it contained part of the internal HR domain. As stated above, the
5   protein between a.a. 168 and 169 (SEQ. ID. No. 23) causes a loss of HR activity.

Because some of the small cloned proteins with 61 a.a. or less were not expressed, several oligopeptides were synthesized with 30 a.a. to narrow down the functional region of the internal HR domain. The oligopeptides were synthesized within the range of a.a. 121-179 (SEQ. ID. No. 23). However, these oligos did not
10   elicit the HR response. It was not expected that there would be an HR response from oligos 137-166, 121-150, and 137-156 (SEQ. ID. No. 23) as these fragments did not contain the imperative amino acids 168 and 169 (SEQ. ID. No. 23). It was expected that the oligo 150-179 (SEQ. ID. No. 23) would elicit an HR response. It is possible that 30 a.a. is too small for the protein to elicit any activity due to a lack of folding
15   and, therefore, a lack of binding or that during the synthesis of the peptides important amino acids were missed (either in the process, or simply by the choice of which 30 amino acids to synthesize) and, therefore, the fragments would not be able to elicit HR. It is also possible, although unlikely, that these small proteins would have undergone some form of post-translational modification within the *E.coli* cell that
20   they did not contain when synthesized and, therefore, were not able to elicit an HR response.

Example 24 – Biological Activity of HR Inducing Fragments

25   The two N-terminal harpin fragments spanning nucleotides 1-104 and nucleotides 1-168 of the nucleic acid of SEQ. ID. No. 24 were effective at inducing resistance of tobacco against TMV, in a similar manner as the full length harpin protein. The internal fragments spanning nucleotides 76-209 and nucleotides 105-209 of the nucleic acid of SEQ. ID. No. 24 were also effective at inducing TMV
30   resistance. In addition, these same four fragments conferred plant growth enhancement ("PGE") in tomato increasing the height of the plants from 4-19% taller than the buffer control plants. The full length harpin protein induced growth

- 55 -

enhancement of 6% greater than the buffer. The negative control did not induce TMV resistance or growth enhancement.

Table 5

5   TMV resistance and PGE activity of HR inducing fragments derived from harpin

Fragment #	Amino acids (SEQ. ID. No. 23)	HR activity	TMV resistance	PGE ht > buffer
1 (+ control)	1-403	+	+	6%
2 (- control)	-	-	-	-2%
9	1-104	+	+	4-8%
10	1-168	+	+	5-13%
13	76-209	+	+	4-18%
15	105-209	+	+	6-19%

Although the invention has been described in detail for the purpose of illustration, it is understood that such detail is solely for that purpose, and variations

10 can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

- 56 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

 (i) APPLICANT: Cornell Research Foundation, Inc. and EDEN Bioscience
   Corporation

 (ii) TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
             FRAGMENTS ELICITING A HYPERSENSITIVE
             RESPONSE AND USES THEREOF

 (iii) NUMBER OF SEQUENCES: 30

 (iv) CORRESPONDENCE ADDRESS:
   (A) ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
   (B) STREET: Clinton Square, P.O. Box 1051
   (C) CITY: Rochester
   (D) STATE: New York
   (E) COUNTRY: U.S.A.
   (F) ZIP: 14603

 (v) COMPUTER READABLE FORM:
   (A) MEDIUM TYPE: Floppy disk
   (B) COMPUTER: IBM PC compatible
   (C) OPERATING SYSTEM: PC-DOS/MS-DOS
   (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

 (vi) CURRENT APPLICATION DATA:
   (A) APPLICATION NUMBER:
   (B) FILING DATE:
   (C) CLASSIFICATION:

 (vii) PRIOR APPLICATION DATA:
   (A) APPLICATION NUMBER: US 60/048,109
   (B) FILING DATE: 30-MAY-1997

 (viii) ATTORNEY/AGENT INFORMATION:
   (A) NAME: Goldman, Michael L.
   (B) REGISTRATION NUMBER: 30,727
   (C) REFERENCE/DOCKET NUMBER: 19603/1302

 (ix) TELECOMMUNICATION INFORMATION:
   (A) TELEPHONE: (716) 263-1304
   (B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

 (i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 31 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

-57-

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGAATTCAT ATGAGTCTGA ATACAAGTGG G   31

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGGAATTCAT ATGGGCGGTG GCTTAGGCGG T   31

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE:.cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCATATGTC GAACGCGCTG AACGATATG   29.

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGAATTCAT ATGTTAGGCG GTTCGCTGAA C   31
(2) INFORMATION FOR SEQ ID NO:5:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 29 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGCATATGCT GAACACGCTG GGCTCGAAA
                            29

(2) INFORMATION FOR SEQ ID NO:6:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 29 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGCATATGTC AACGTCCCAA AACGACGAT
                            29

(2) INFORMATION FOR SEQ ID NO:7:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 27 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGCATATGTC CACCTCAGAC TCCAGCG
                            27

(2) INFORMATION FOR SEQ ID NO:8:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 34 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGAATTCAT ATGCAAAGCC TGTTTGGTGA TGGG   34

(2) INFORMATION FOR SEQ ID NO:9:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGAATTCAT ATGGGTAATG GTCTGAGCAA G   31

(2) INFORMATION FOR SEQ ID NO:10:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGAATTCAT ATGAAAGCGG GCATTCAGGC G   31

(2) INFORMATION FOR SEQ ID NO:11:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 34 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
GGGAATTCAT ATGACACCAG CCAGTATGGA GCAG   34

- 60 -

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCAAGCTTAA CAGCCCACCA CCGCCCATCA T   31

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCAAGCTTAA ATCGTTCAGC GCGTTCGACA G   31

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 34 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCAAGCTTAA TATCTCGCTG AACATCTTCA GCAG   34

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCAAGCTTAA GGTGCCATCT TGCCCATCAC   30

    (2) INFORMATION FOR SEQ ID NO:16:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 34 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAAGCTTAA ATCAGTGACT CCTTTTTTAT AGGC   34

    (2) INFORMATION FOR SEQ ID NO:17:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCAAGCTTAA CAGGCCCGAC AGCGCATCAG T   31

    (2) INFORMATION FOR SEQ ID NO:18:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCAAGCTTAA ACCGATACCG GTACCCACGG C   31

- 62 -

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 34 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCAAGCTTAA TCCGTCGTCA TCTGGCTTGC TCAG   34

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCAAGCTTAA GCCGCGCCCA GCTTG   25

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 338 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Gln Ile Thr Ile Lys Ala His Ile Gly Gly Asp Leu Gly Val Ser
1   5   10   15

Gly Leu Gly Ala Gln Gly Leu Lys Gly Leu Asn Ser Ala Ala Ser Ser
            20   25   30

Leu Gly Ser Ser Val Asp Lys Leu Ser Ser Thr Ile Asp Lys Leu Thr
        35   40   45

```
Ser Ala Leu Thr Ser Met Met Phe Gly Gly Ala Leu Ala Gln Gly Leu
    50                  55              60

Gly Ala Ser Ser Lys Gly Leu Gly Met Ser Asn Gln Leu Gly Gln Ser
65                  70              75                  80

Phe Gly Asn Gly Ala Gln Gly Ala Ser Asn Leu Leu Ser Val Pro Lys
            85              90                  95

Ser Gly Gly Asp Ala Leu Ser Lys Met Phe Asp Lys Ala Leu Asp Asp
            100             105             110

Leu Leu Gly His Asp Thr Val Thr Lys Leu Thr Asn Gln Ser Asn Gln
        115             120             125

Leu Ala Asn Ser Met Leu Asn Ala Ser Gln Met Thr Gln Gly Asn Met
    130             135             140

Asn Ala Phe Gly Ser Gly Val Asn Asn Ala Leu Ser Ser Ile Leu Gly
145             150             155             160

Asn Gly Leu Gly Gln Ser Met Ser Gly Phe Ser Gln Pro Ser Leu Gly
                165                 170                 175

Ala Gly Gly Leu Gln Gly Leu Ser Gly Ala Gly Ala Phe Asn Gln Leu
            180             185             190

Gly Asn Ala Ile Gly Met Gly Val Gly Gln Asn Ala Ala Leu Ser Ala
        195             200             205

Leu Ser Asn Val Ser Thr His Val Asp Gly Asn Asn Arg His Phe Val
    210             215             220

Asp Lys Glu Asp Arg Gly Met Ala Lys Glu Ile Gly Gln Phe Met Asp
225             230             235             240

Gln Tyr Pro Glu Ile Phe Gly Lys Pro Glu Tyr Gln Lys Asp Gly Trp
            245             250             255

Ser Ser Pro Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser Lys
            260             265             270

Pro Asp Asp Asp Gly Met Thr Gly Ala Ser Met Asp Lys Phe Arg Gln
        275             280             285

Ala Met Gly Met Ile Lys Ser Ala Val Ala Gly Asp Thr Gly Asn Thr
    290             295             300

Asn Leu Asn Leu Arg Gly Ala Gly Gly Ala Ser Leu Gly Ile Asp Ala
305             310             315             320

Ala Val Val Gly Asp Lys Ile Ala Asn Met Ser Leu Gly Lys Leu Ala
            325             330             335

Asn Ala
```

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2141 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
CGATTTTACC CGGGTGAACG TGCTATGACC GACAGCATCA CGGTATTCGA CACCGTTACG      60

GCGTTTATGG CCGCGATGAA CCGGCATCAG GCGGCGCGCT GGTCGCCGCA ATCCGGCGTC     120

GATCTGGTAT TTCAGTTTGG GGACACCGGG CGTGAACTCA TGATGCAGAT TCAGCCGGGG     180

CAGCAATATC CCGGCATGTT GCGCACGCTG CTCGCTCGTC GTTATCAGCA GGCGGCAGAG     240

TGCGATGGCT GCCATCTGTG CCTGAACGGC AGCGATGTAT TGATCCTCTG GTGGCCGCTG     300

CCGTCGGATC CCGGCAGTTA TCCGCAGGTG ATCGAACGTT TGTTTGAACT GGCGGGAATG     360

ACGTTGCCGT CGCTATCCAT AGCACCGACG GCGCGTCCGC AGACAGGGAA CGGACGCGCC     420

CGATCATTAA GATAAAGGCG GCTTTTTTTA TTGCAAAACG GTAACGGTGA GGAACCGTTT     480

CACCGTCGGC GTCACTCAGT AACAAGTATC CATCATGATG CCTACATCGG GATCGGCGTG     540

GGCATCCGTT GCAGATACTT TTGCGAACAC CTGACATGAA TGAGGAAACG AAATTATGCA     600

AATTACGATC AAAGCGCACA TCGGCGGTGA TTTGGGCGTC TCCGGTCTGG GGCTGGGTGC     660

TCAGGGACTG AAAGGACTGA ATTCCGCGGC TTCATCGCTG GGTTCCAGCG TGGATAAACT     720

GAGCAGCACC ATCGATAAGT TGACCTCCGC GCTGACTTCG ATGATGTTTG GCGGCGCGCT     780

GGCGCAGGGG CTGGGCGCCA GCTCGAAGGG GCTGGGGATG AGCAATCAAC TGGGCCAGTC     840

TTTCGGCAAT GGCGCGCAGG GTGCGAGCAA CCTGCTATCC GTACCGAAAT CCGGCGGCGA     900

TGCGTTGTCA AAAATGTTTG ATAAAGCGCT GGACGATCTG CTGGGTCATG ACACCGTGAC     960

CAAGCTGACT AACCAGAGCA ACCAACTGGC TAATTCAATG CTGAACGCCA GCCAGATGAC    1020

CCAGGGTAAT ATGAATGCGT TCGGCAGCGG TGTGAACAAC GCACTGTCGT CCATTCTCGG    1080

CAACGGTCTC GGCCAGTCGA TGAGTGGCTT CTCTCAGCCT TCTCTGGGGG CAGGCGGCTT    1140

GCAGGGCCTG AGCGGCGCGG GTGCATTCAA CCAGTTGGGT AATGCCATCG GCATGGGCGT    1200

GGGGCAGAAT GCTGCGCTGA GTGCGTTGAG TAACGTCAGC ACCCACGTAG ACGGTAACAA    1260
```

```
CCGCCACTTT GTAGATAAAG AAGATCGCGG CATGGCGAAA GAGATCGGCC AGTTTATGGA    1320

TCAGTATCCG GAAATATTCG GTAAACCGGA ATACCAGAAA GATGGCTGGA GTTCGCCGAA    1380

GACGGACGAC AAATCCTGGG CTAAAGCGCT GAGTAAACCG GATGATGACG GTATGACCGG    1440

CGCCAGCATG GACAAATTCC GTCAGGCGAT GGGTATGATC AAAAGCGCGG TGGCGGGTGA    1500

TACCGGCAAT ACCAACCTGA ACCTGCGTGG CGCGGGCGGT GCATCGCTGG GTATCGATGC    1560

GGCTGTCGTC GGCGATAAAA TAGCCAACAT GTCGCTGGGT AAGCTGGCCA ACGCCTGATA    1620

ATCTGTGCTG GCCTGATAAA GCGGAAACGA AAAAGAGAC GGGGAAGCCT GTCTCTTTTC     1680

TTATTATGCG GTTTATGCGG TTACCTGGAC CGGTTAATCA TCGTCATCGA TCTGGTACAA    1740

ACGCACATTT TCCCGTTCAT TCGCGTCGTT ACGCGCCACA ATCGCGATGG CATCTTCCTC    1800

GTCGCTCAGA TTGCGCGGCT GATGGGGAAC GCCGGGTGGA ATATAGAGAA ACTCGCCGGC    1860

CAGATGGAGA CACGTCTGCG ATAAATCTGT GCCGTAACGT GTTTCTATCC GCCCCTTTAG    1920

CAGATAGATT GCGGTTTCGT AATCAACATG GTAATGCGGT TCCGCCTGTG CGCCGGCCGG    1980

GATCACCACA ATATTCATAG AAAGCTGTCT TGCACCTACC GTATCGCGGG AGATACCGAC    2040

AAAATAGGGC AGTTTTTGCG TGGTATCCGT GGGGTGTTCC GGCCTGACAA TCTTGAGTTG    2100

GTTCGTCATC ATCTTTCTCC ATCTGGGCGA CCTGATCGGT T                        2141
```

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 403 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Met Ser Leu Asn Thr Ser Gly Leu Gly Ala Ser Thr Met Gln Ile Ser
1               5                   10                  15

Ile Gly Gly Ala Gly Gly Asn Asn Gly Leu Leu Gly Thr Ser Arg Gln
            20                  25                  30

Asn Ala Gly Leu Gly Gly Asn Ser Ala Leu Gly Leu Gly Gly Gly Asn
        35                  40                  45
```

```
Gln Asn Asp Thr Val Asn Gln Leu Ala Gly Leu Leu Thr Gly Met Met
    50              55              60

Met Met Met Ser Met Met Gly Gly Gly Gly Leu Met Gly Gly Gly Leu
65              70              75              80

Gly Gly Gly Leu Gly Asn Gly Leu Gly Gly Ser Gly Gly Leu Gly Glu
            85              90              95

Gly Leu Ser Asn Ala Leu Asn Asp Met Leu Gly Gly Ser Leu Asn Thr
            100             105             110

Leu Gly Ser Lys Gly Gly Asn Asn Thr Thr Ser Thr Thr Asn Ser Pro
    115             120             125

Leu Asp Gln Ala Leu Gly Ile Asn Ser Thr Ser Gln Asn Asp Asp Ser
    130             135             140

Thr Ser Gly Thr Asp Ser Thr Ser Asp Ser Ser Asp Pro Met Gln Gln
145             150             155             160

Leu Leu Lys Met Phe Ser Glu Ile Met Gln Ser Leu Phe Gly Asp Gly
            165             170             175

Gln Asp Gly Thr Gln Gly Ser Ser Ser Gly Gly Lys Gln Pro Thr Glu
            180             185             190

Gly Glu Gln Asn Ala Tyr Lys Lys Gly Val Thr Asp Ala Leu Ser Gly
    195             200             205

Leu Met Gly Asn Gly Leu Ser Gln Leu Leu Gly Asn Gly Gly Leu Gly
    210             215             220

Gly Gly Gln Gly Gly Asn Ala Gly Thr Gly Leu Asp Gly Ser Ser Leu
225             230             235             240

Gly Gly Lys Gly Leu Gln Asn Leu Ser Gly Pro Val Asp Tyr Gln Gln
            245             250             255

Leu Gly Asn Ala Val Gly Thr Gly Ile Gly Met Lys Ala Gly Ile Gln
            260             265             270

Ala Leu Asn Asp Ile Gly Thr His Arg His Ser Ser Thr Arg Ser Phe
    275             280             285

Val Asn Lys Gly Asp Arg Ala Met Ala Lys Glu Ile Gly Gln Phe Met
    290             295             300

Asp Gln Tyr Pro Glu Val Phe Gly Lys Pro Gln Tyr Gln Lys Gly Pro
305             310             315             320

Gly Gln Glu Val Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser
            325             330             335
```

- 67 -

```
        Lys Pro Asp Asp Asp Gly Met Thr Pro Ala Ser Met Glu Gln Phe Asn
                340                     345             350

        Lys Ala Lys Gly Met Ile Lys Arg Pro Met Ala Gly Asp Thr Gly Asn
                355                     360             365

        Gly Asn Leu Gln Ala Arg Gly Ala Gly Gly Ser Ser Leu Gly Ile Asp
                370             375             380

        Ala Met Met Ala Gly Asp Ala Ile Asn Asn Met Ala Leu Gly Lys Leu
        385             · 390             395             400

        Gly Ala Ala
```

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1288 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
AAGCTTCGGC ATGGCACGTT TGACCGTTGG GTCGGCAGGG TACGTTTGAA TTATTCATAA      60

GAGGAATACG TTATGAGTCT GAATACAAGT GGGCTGGGAG CGTCAACGAT GCAAATTTCT     120

ATCGGCGGTG CGGGCGGAAA TAACGGGTTG CTGGGTACCA GTCGCCAGAA TGCTGGGTTG     180

GGTGGCAATT CTGCACTGGG GCTGGGCGGC GGTAATCAAA ATGATACCGT CAATCAGCTG     240

GCTGGCTTAC TCACCGGCAT GATGATGATG ATGAGCATGA TGGGCGGTGG TGGGCTGATG     300.

GGCGGTGGCT TAGGCGGTGG CTTAGGTAAT GGCTTGGGTG GCTCAGGTGG CCTGGGCGAA     360

GGACTGTCGA ACGCGCTGAA CGATATGTTA GGCGGTTCGC TGAACACGCT GGGCTCGAAA     420

GGCGGCAACA ATACCACTTC AACAACAAAT TCCCCGCTGG ACCAGGCGCT GGGTATTAAC     480

TCAACGTCCC AAAACGACGA TTCCACCTCC GGCACAGATT CCACCTCAGA CTCCAGCGAC     540

CCGATGCAGC AGCTGCTGAA GATGTTCAGC GAGATAATGC AAAGCCTGTT TGGTGATGGG     600

CAAGATGGCA CCCAGGGCAG TTCCTCTGGG GGCAAGCAGC CGACCGAAGG CGAGCAGAAC     660

GCCTATAAAA AAGGAGTCAC TGATGCGCTG TCGGGCCTGA TGGGTAATGG TCTGAGCCAG     720

CTCCTTGGCA ACGGGGGACT GGGAGGTGGT CAGGGCGGTA ATGCTGGCAC GGGTCTTGAC     780
```

- 68 -

GGTTCGTCGC TGGGCGGCAA AGGGCTGCAA AACCTGAGCG GGCCGGTGGA CTACCAGCAG   840

TTAGGTAACG CCGTGGGTAC CGGTATCGGT ATGAAAGCGG GCATTCAGGC GCTGAATGAT   900

ATCGGTACGC ACAGGCACAG TTCAACCCGT TCTTTCGTCA ATAAAGGCGA TCGGGCGATG   960

GCGAAGGAAA TCGGTCAGTT CATGGACCAG TATCCTGAGG TGTTTGGCAA GCCGCAGTAC   1020

CAGAAAGGCC CGGGTCAGGA GGTGAAAACC GATGACAAAT CATGGGCAAA AGCACTGAGC   1080

AAGCCAGATG ACGACGGAAT GACACCAGCC AGTATGGAGC AGTTCAACAA AGCCAAGGGC   1140

ATGATCAAAA GGCCCATGGC GGGTGATACC GGCAACGGCA ACCTGCAGGC ACGCGGTGCC   1200

GGTGGTTCTT CGCTGGGTAT TGATGCCATG ATGGCCGGTG ATGCCATTAA CAATATGGCA   1260

CTTGGCAAGC TGGGCGCGGC TTAAGCTT   1288

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 341 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS:
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met Gln Ser Leu Ser Leu Asn Ser Ser Ser Leu Gln Thr Pro Ala Met
1   5   10   15

Ala Leu Val Leu Val Arg Pro Glu Ala Glu Thr Thr Gly Ser Thr Ser
            20   25   30

Ser Lys Ala Leu Gln Glu Val Val Val Lys Leu Ala Glu Glu Leu Met
        35   40   45

Arg Asn Gly Gln Leu Asp Asp Ser Ser Pro Leu Gly Lys Leu Leu Ala
        50   55   60

Lys Ser Met Ala Ala Asp Gly Lys Ala Gly Gly Gly Ile Glu Asp Val
65   70   75   80

Ile Ala Ala Leu Asp Lys Leu Ile His Glu Lys Leu Gly Asp Asn Phe
                85   90   95

Gly Ala Ser Ala Asp Ser Ala Ser Gly Thr Gly Gln Gln Asp Leu Met
            100   105   110

Thr Gln Val Leu Asn Gly Leu Ala Lys Ser Met Leu Asp Asp Leu Leu
            115   120   125

- 69 -

```
Thr Lys Gln Asp Gly Gly Thr Ser Phe Ser Glu Asp Asp Met Pro Met
    130             135             140

Leu Asn Lys Ile Ala Gln Phe Met Asp Asp Asn Pro Ala Gln Phe Pro
145             150             155             160

Lys Pro Asp Ser Gly Ser Trp Val Asn Glu Leu Lys Glu Asp Asn Phe
            165             170             175

Leu Asp Gly Asp Glu Thr Ala Ala Phe Arg Ser Ala Leu Asp Ile Ile
            180             185             190

Gly Gln Gln Leu Gly Asn Gln Gln Ser Asp Ala Gly Ser Leu Ala Gly
        195             200             205

Thr Gly Gly Leu Gly Thr Pro Ser Ser Phe Ser Asn Asn Ser Ser
    210             215             220

Val Met Gly Asp Pro Leu Ile Asp Ala Asn Thr Gly Pro Gly Asp Ser
225             230             235             240

Gly Asn Thr Arg Gly Glu Ala Gly Gln Leu Ile Gly Glu Leu Ile Asp
            245             250             255

Arg Gly Leu Gln Ser Val Leu Ala Gly Gly Gly Leu Gly Thr Pro Val
            260             265             270

Asn Thr Pro Gln Thr Gly Thr Ser Ala Asn Gly Gly Gln Ser Ala Gln
        275             280             285

Asp Leu Asp Gln Leu Leu Gly Gly Leu Leu Leu Lys Gly Leu Glu Ala
    290             295             300

Thr Leu Lys Asp Ala Gly Gln Thr Gly Thr Asp Val Gln Ser Ser Ala
305             310             315             320

Ala Gln Ile Ala Thr Leu Leu Val Ser Thr Leu Leu Gln Gly Thr Arg
            325             330             335

Asn Gln Ala Ala Ala
        340
```

(2) INFORMATION FOR SEQ ID NO:26:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 1026 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

- 70 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
ATGCAGAGTC TCAGTCTTAA CAGCAGCTCG CTGCAAACCC CGGCAATGGC CCTTGTCCTG      60

GTACGTCCTG AAGCCGAGAC GACTGGCAGT ACGTCGAGCA AGGCGCTTCA GGAAGTTGTC     120

GTGAAGCTGG CCGAGGAACT GATGCGCAAT GGTCAACTCG ACGACAGCTC GCCATTGGGA     180

AAACTGTTGG CCAAGTCGAT GGCCGCAGAT GGCAAGGCGG GCGGCGGTAT TGAGGATGTC     240

ATCGCTGCGC TGGACAAGCT GATCCATGAA AAGCTCGGTG ACAACTTCGG CGCGTCTGCG     300

GACAGCGCCT CGGGTACCGG ACAGCAGGAC CTGATGACTC AGGTGCTCAA TGGCCTGGCC     360

AAGTCGATGC TCGATGATCT TCTGACCAAG CAGGATGGCG GGACAAGCTT CTCCGAAGAC     420

GATATGCCGA TGCTGAACAA GATCGCGCAG TTCATGGATG ACAATCCCGC ACAGTTTCCC     480

AAGCCGGACT CGGGCTCCTG GGTGAACGAA CTCAAGGAAG ACAACTTCCT TGATGGCGAC     540

GAAACGGCTG CGTTCCGTTC GGCACTCGAC ATCATTGGCC AGCAACTGGG TAATCAGCAG     600

AGTGACGCTG GCAGTCTGGC AGGGACGGGT GGAGGTCTGG GCACTCCGAG CAGTTTTTCC     660

AACAACTCGT CCGTGATGGG TGATCCGCTG ATCGACGCCA ATACCGGTCC CGGTGACAGC     720

GGCAATACCC GTGGTGAAGC GGGGCAACTG ATCGGCGAGC TTATCGACCG TGGCCTGCAA     780

TCGGTATTGG CCGGTGGTGG ACTGGGCACA CCCGTAAACA CCCCGCAGAC CGGTACGTCG     840

GCGAATGGCG GACAGTCCGC TCAGGATCTT GATCAGTTGC TGGGCGGCTT GCTGCTCAAG     900

GGCCTGGAGG CAACGCTCAA GGATGCCGGG CAAACAGGCA CCGACGTGCA GTCGAGCGCT     960

GCGCAAATCG CCACCTTGCT GGTCAGTACG CTGCTGCAAG GCACCCGCAA TCAGGCTGCA    1020

GCCTGA                                                                1026
```

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 344 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Met Ser Val Gly Asn Ile Gln Ser Pro Ser Asn Leu Pro Gly Leu Gln
1               5                   10                  15
```

- 71 -

```
Asn Leu Asn Leu Asn Thr Asn Thr Asn Ser Gln Gln Ser Gly Gln Ser
            20              25              30

Val Gln Asp Leu Ile Lys Gln Val Glu Lys Asp Ile Leu Asn Ile Ile
            35              40              45

Ala Ala Leu Val Gln Lys Ala Ala Gln Ser Ala Gly Gly Asn Thr Gly
    50              55              60

Asn Thr Gly Asn Ala Pro Ala Lys Asp Gly Asn Ala Asn Ala Gly Ala
65              70              75              80

Asn Asp Pro Ser Lys Asn Asp Pro Ser Lys Ser Gln Ala Pro Gln Ser
            85              90              95

Ala Asn Lys Thr Gly Asn Val Asp Asp Ala Asn Asn Gln Asp Pro Met
            100             105             110

Gln Ala Leu Met Gln Leu Leu Glu Asp Leu Val Lys Leu Leu Lys Ala
            115             120             125

Ala Leu His Met Gln Gln Pro Gly Gly Asn Asp Lys Gly Asn Gly Val
    130             135             140

Gly Gly Ala Asn Gly Ala Lys Gly Ala Gly Gly Gln Gly Gly Leu Ala
145             150             155             160

Glu Ala Leu Gln Glu Ile Glu Gln Ile Leu Ala Gln Leu Gly Gly Gly
            165             170             175

Gly Ala Gly Ala Gly Gly Ala Gly Gly Gly Val Gly Gly Ala Gly Gly
            180             185             190

Ala Asp Gly Gly Ser Gly Ala Gly Gly Ala Gly Gly Ala Asn Gly Ala
            195             200             205

Asp Gly Gly Asn Gly Val Asn Gly Asn Gln Ala Asn Gly Pro Gln Asn
    210             215             220

Ala Gly Asp Val Asn Gly Ala Asn Gly Ala Asp Asp Gly Ser Glu Asp
225             230             235             240

Gln Gly Gly Leu Thr Gly Val Leu Gln Lys Leu Met Lys Ile Leu Asn
            245             250             255

Ala Leu Val Gln Met Met Gln Gln Gly Gly Leu Gly Gly Gly Asn Gln
            260             265             270

Ala Gln Gly Gly Ser Lys Gly Ala Gly Asn Ala Ser Pro Ala Ser Gly
    275             280             285

Ala Asn Pro Gly Ala Asn Gln Pro Gly Ser Ala Asp Asp Gln Ser Ser
    290             295             300
```

- 72 -

```
Gly Gln Asn Asn Leu Gln Ser Gln Ile Met Asp Val Val Lys Glu Val
305               310               315               320

Val Gln Ile Leu Gln Gln Met Leu Ala Ala Gln Asn Gly Gly Ser Gln
              325               330               335

Gln Ser Thr Ser Thr Gln Pro Met
              340
```

(2) INFORMATION FOR SEQ ID NO:28:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1035 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
ATGTCAGTCG GAAACATCCA GAGCCCGTCG AACCTCCCGG GTCTGCAGAA CCTGAACCTC      60

AACACCAACA CCAACAGCCA GCAATCGGGC CAGTCCGTGC AAGACCTGAT CAAGCAGGTC     120

GAGAAGGACA TCCTCAACAT CATCGCAGCC CTCGTGCAGA AGGCCGCACA GTCGGCGGGC     180

GGCAACACCG GTAACACCGG CAACGCGCCG GCGAAGGACG GCAATGCCAA CGCGGGCGCC     240

AACGACCCGA GCAAGAACGA CCCGAGCAAG AGCCAGGCTC CGCAGTCGGC CAACAAGACC     300

GGCAACGTCG ACGACGCCAA CAACCAGGAT CCGATGCAAG CGCTGATGCA GCTGCTGGAA     360

GACCTGGTGA AGCTGCTGAA GGCGGCCCTG CACATGCAGC AGCCCGGCGG CAATGACAAG     420

GGCAACGGCG TGGGCGGTGC CAACGGCGCC AAGGGTGCCG GCGGCCAGGG CGGCCTGGCC     480

GAAGCGCTGC AGGAGATCGA GCAGATCCTC GCCCAGCTCG GCGGCGGCGG TGCTGGCGCC     540

GGCGGCGCGG GTGGCGGTGT CGGCGGTGCT GGTGGCGCGG ATGGCGGCTC CGGTGCGGGT     600

GGCGCAGGCG GTGCGAACGG CGCCGACGGC GGCAATGGCG TGAACGGCAA CCAGGCGAAC     660

GGCCCGCAGA ACGCAGGCGA TGTCAACGGT GCCAACGGCG CGGATGACGG CAGCGAAGAC     720

CAGGGCGGCC TCACCGGCGT GCTGCAAAAG CTGATGAAGA TCCTGAACGC GCTGGTGCAG     780

ATGATGCAGC AAGGCGGCCT CGGCGGCGGC AACCAGGCGC AGGGCGGCTC GAAGGGTGCC     840

GGCAACGCCT CGCCGGCTTC CGGCGCGAAC CCGGGCGCGA ACCAGCCCGG TTCGGCGGAT     900

GATCAATCGT CCGGCCAGAA CAATCTGCAA TCCCAGATCA TGGATGTGGT GAAGGAGGTC     960
```

- 73 -

GTCCAGATCC TGCAGCAGAT GCTGGCGGCG CAGAACGGCG GCAGCCAGCA GTCCACCTCG   1020

ACGCAGCCGA TGTAA   1035

(2) INFORMATION FOR SEQ ID NO:29:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Thr Leu Ile Glu Leu Met Ile Val Val Ala Ile Ile Ala Ile Leu Ala
1   5   10   15

Ala Ile Ala Leu Pro Ala Tyr Gln Asp Tyr
            20   25

(2) INFORMATION FOR SEQ ID NO:30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ser Ser Gln Gln Ser Pro Ser Ala Gly Ser Glu Gln Gln Leu Asp Gln
1   5   10   15

Leu Leu Ala Met
            20

- 74 -

WHAT IS CLAIMED:

1.   An isolated fragment of an *Erwinia* hypersensitive response elicitor protein or polypeptide, wherein said fragment elicits a hypersensitive response in

5   plants.

2.   An isolated fragment according to claim 1, wherein the hypersensitive response elicitor protein or polypeptide is derived from an *Erwinia amylovora, Erwinia carotovora, Erwinia chrysanthemi, or Erwinia stewartii.*

10

3.   An isolated fragment according to claim 2, wherein the hypersensitive response elicitor protein or polypeptide is derived from *Erwinia amylovora.*

4.   An isolated fragment according to claim 3, wherein the fragment is

15   selected from the group consisting of a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, and an internal fragment of the amino acid sequence of SEQ. ID. No. 23.

20   5.   An isolated fragment according to claim 4, wherein the fragment is a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning amino acids 105 and 403 of SEQ. ID. No. 23.

6.   An isolated fragment according to claim 4, wherein the fragment is an

25   N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 1 and 98, 1 and 104, 1 and 122, 1 and 168, 1 and 218, 1 and 266, 1 and 342, 1 and 321, and 1 and 372.

7.   An isolated fragment according to claim 4, wherein the fragment is an

30   internal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 76 and 209, 105 and 209, 99 and 209, 137 and 204, 137 and 200, 109 and 204, 109 and 200, 137 and 180, and 105 and 180.

- 75 -

8.   An isolated DNA molecule encoding a fragment according to claim 1.

9.   An isolated DNA molecule according to claim 8, wherein the hypersensitive response elicitor protein or polypeptide is derived from an *Erwinia*

5   *amylovora, Erwinia carotovora, Erwinia chrysanthemi, or Erwinia stewartii.*

10.   An isolated DNA molecule according to claim 9, wherein the hypersensitive response elicitor protein or polypeptide is derived from *Erwinia amylovora.*

10

11.   An isolated isolated DNA molecule according to claim 10, wherein the fragment is selected from the group consisting of a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, and an internal fragment of the amino acid sequence of

15   SEQ. ID. No. 23.

12.   An isolated DNA molecule according to claim 10, wherein the fragment is a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning amino acids 105 and 403 of SEQ. ID. No. 23.

20

13.   An isolated DNA molecule according to claim 10, wherein the fragment is an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 1 and 98, 1 and 104, 1 and 122, 1 and 168, 1 and 218, 1 and 266, 1 and 342, 1 and 321, and 1 and 372.

25

14.   An isolated DNA molecule according to claim 10, wherein the fragment is an internal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 76 and 209, 105 and 209, 99 and 209, 137 and 204, 137 and 200, 109 and 204, 109 and 200, 137 and 180, and 105

30   and 180.

15. An expression system transformed with a DNA molecule according to claim 8.

16. An expression system according to claim 15, wherein said DNA
5   molecule is in proper sense orientation and correct reading frame.

17. A host cell transformed with a DNA molecule according to claim 8.

18. A host cell according to claim 17, wherein the host cell is selected
10   from the group consisting of a plant cell and a bacterial cell.

19. A host cell according to claim 17, wherein the DNA molecule is transformed with an expression system.

15   20. A transgenic plant transformed with the DNA molecule of claim 8.

21. A transgenic plant according to claim 20, wherein the plant is selected from the group consisting of alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage,
20   brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane.

25   22. A transgenic plant according to claim 20, wherein the plant is selected from the group consisting of *Arabidopsis thaliana*, *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

23. A transgenic plant seed transformed with the DNA molecule of
30   claim 8.

- 77 -

24.   A transgenic plant seed according to claim 23, wherein the plant seed
is selected from the group consisting of alfalfa, rice, wheat, barley, rye, cotton,
sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive,
cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish,

5   spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,
cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple,
soybean, tobacco, tomato, sorghum, and sugarcane.

25.   A transgenic plant seed according to claim 23, wherein the plant seed

10   is selected from the group consisting of *Arabidopsis thaliana*, *Saintpaulia*, petunia,
pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

26.   A method of imparting disease resistance to plants comprising:
applying a fragment of a hypersensitive response elicitor protein or

15   polypeptide, which fragment elicits a hypersensitive response, in a non-infectious
form to a plant or plant seed under conditions impart disease resistance.

27.   A method according to claim 26, wherein plants are treated during said
applying.

20

28.   A method according to claim 26, wherein plant seeds are treated during
said applying, said method further comprising:
planting the seeds treated with the fragment of the hypersensitive
response elicitor in natural or artificial soil and

25   propagating plants from the seeds planted in the soil.

29.   A method of enhancing plant growth comprising:
applying a fragment of a hypersensitive response elicitor protein or
polypeptide, which fragment elicits a hypersensitive response, in a non-infectious

30   form to a plant or plant seed under conditions effective to enhance plant growth.
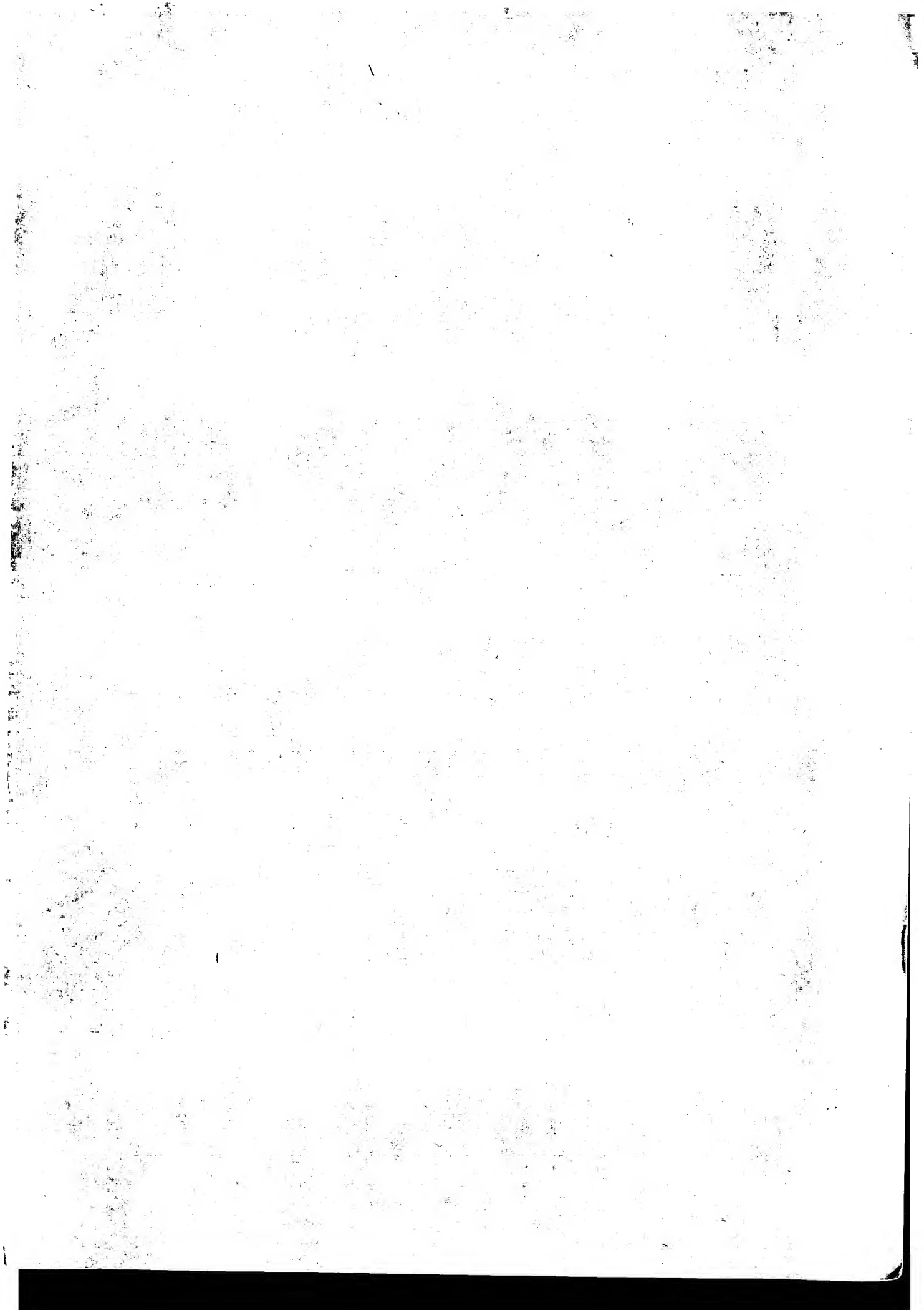
30.   A method according to claim 29, wherein plants are treated during said applying.

31.   A method according to claim 29, wherein plant seeds are treated during said applying, said method further comprising:

planting the seeds treated with the fragment of the hypersensitive response elicitor in natural or artificial soil and

propagating plants from the seeds planted in the soil.

32.   A method of insect control for plants comprising:

applying a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, in a non-infectious form to a plant or plant seed under conditions effective to control insects.

33.   A method according to claim 32, wherein plants are treated during said applying.

34.   A method according to claim 32, wherein plant seeds are treated during said applying, said method further comprising:

planting the seeds treated with the fragment of the hypersensitive response elicitor in natural or artificial soil and

propagating plants from the seeds planted in the soil.

35.   A method of imparting disease resistance to plants comprising:

.   providing a transgenic plant or plant seed transformed with a DNA molecule which encodes a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, and

growing the transgenic plant or transgenic plants produced from the transgenic plant seeds under conditions effective to impart disease resistance.

36.   A method according to claim 35, wherein a transgenic plant is provided.

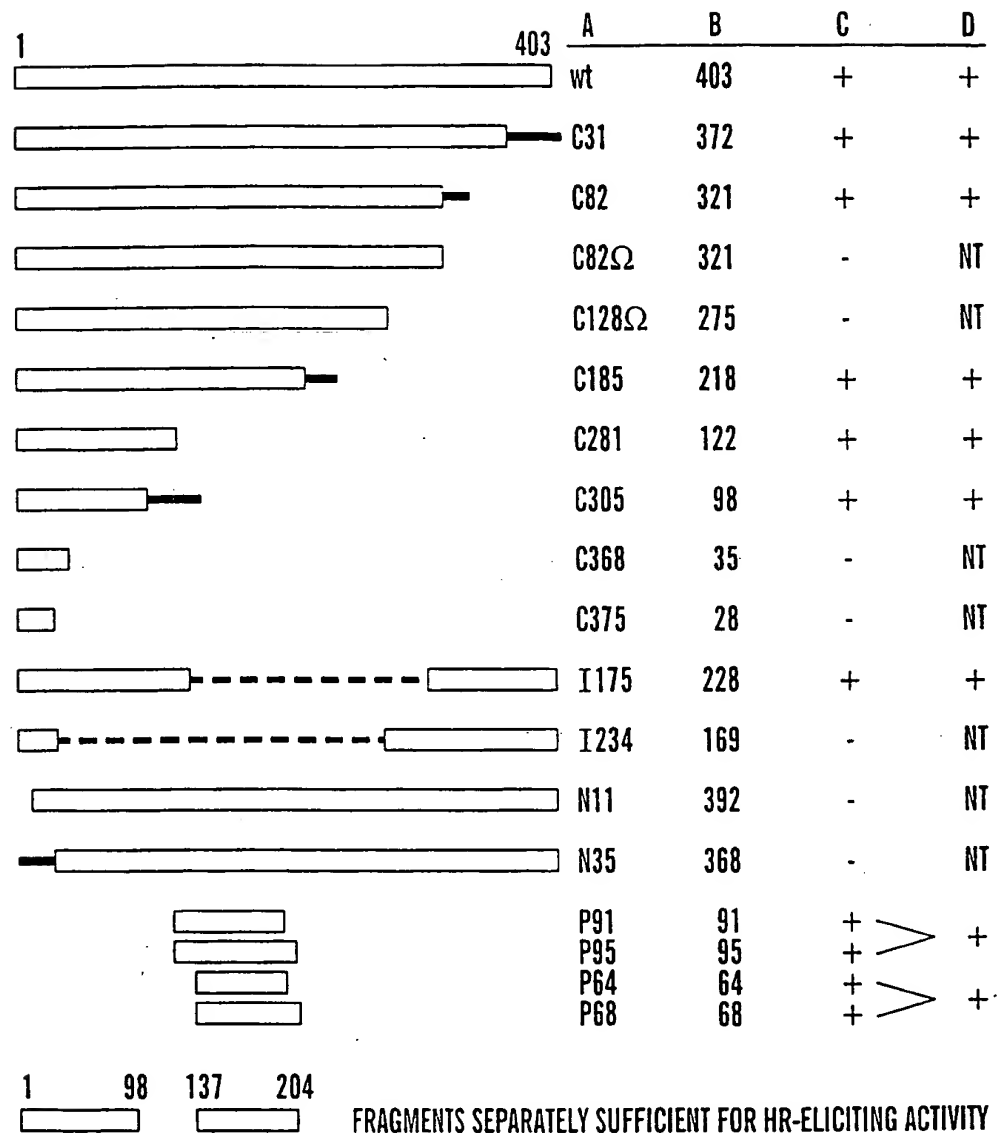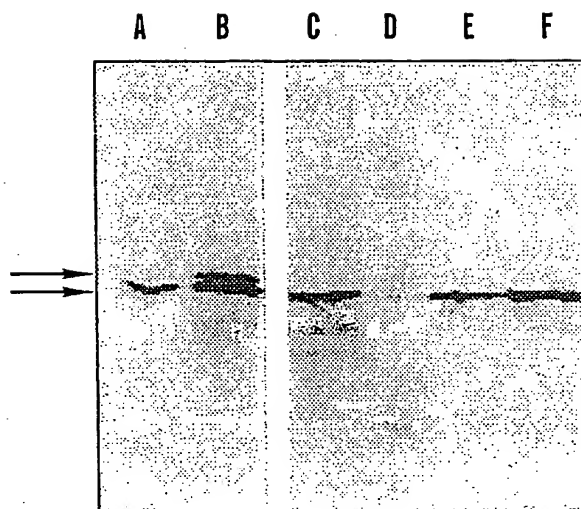37.   A method according to claim 35, wherein a transgenic plant seed is provided.
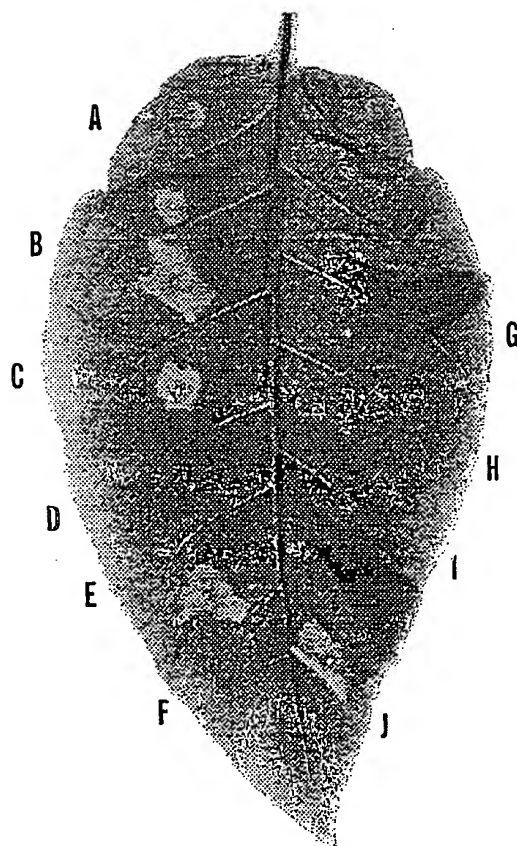
5   38.   A method of enhancing plant growth comprising:

providing a transgenic plant or a plant seed transformed with a DNA molecule which encodes a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, and

growing the transgenic plant or transgenic plants produced from the

10   transgenic plant seeds under conditions effective to enhance plant growth.

39.   A method according to claim 38, wherein a transgenic plant is provided.

15   40.   A method according to claim 38, wherein a transgenic plant seed is provided.

41.   A method of insect control for plants comprising:

providing a transgenic plant or plant seed transformed with a DNA

20   molecule which encodes a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, and

growing the transgenic plant or transgenic plants produced from the transgenic plant seeds under conditions effective to control insects.

25   42.   A method according to claim 41, wherein a transgenic plant is provided.

43.   A method according to claim 41, wherein a transgenic plant seed is provided.

30

1/13

	A	B	C	D
wt		403	+	+
	C31	372	+	+
	C82	321	+	+
	C82Ω	321	-	NT
	C128Ω	275	-	NT
	C185	218	+	+
	C281	122	+	+
	C305	98	+	+
	C368	35	-	NT
	C375	28	-	NT
	I175	228	+	+
	I234	169	-	NT
	N11	392	-	NT
	N35	368	-	NT
	P91	91	+	+
	P95	95	+	+
	P64	64	+	+
	P68	68	+	+

1   98   137   204

FRAGMENTS SEPARATELY SUFFICIENT FOR HR-ELICITING ACTIVITY

FIG. 1

FIG. 2

FIG. 3

FIG. 4

FIG. 5A

FIG. 5B

FIG. 5C

FIG. 6

SMEP (METALLOPROTEINASE) CLEAVAGE SITES

ENDOPROTEINASE Glu-C CLEAVAGE SITES (Glu)

TRYPSIN CLEAVAGE SITES

ACTIVE FRAGMENTS

↓ = CLEAVAGE AT THESE SITES ABOLISHES HR-ELICITING ACTIVITY OF HARPIN C305 AND OF PEPTIDES P64/P68 AND P91/P95
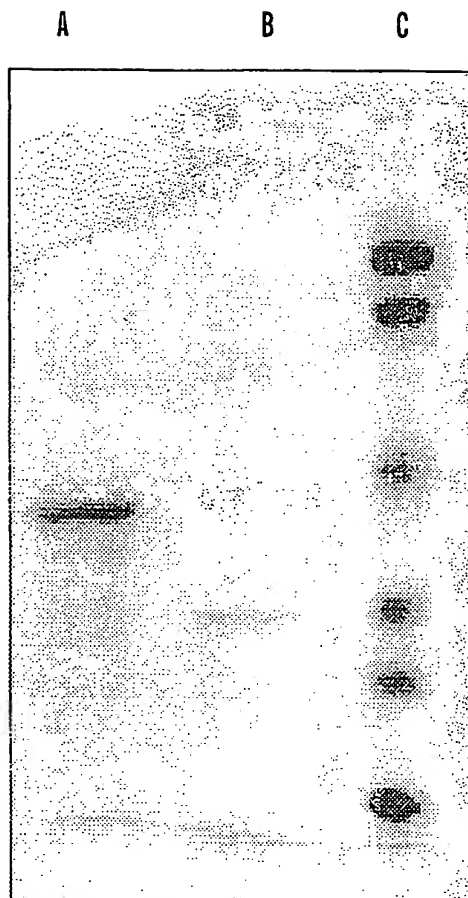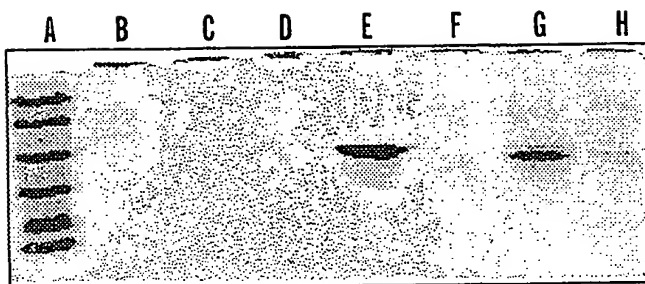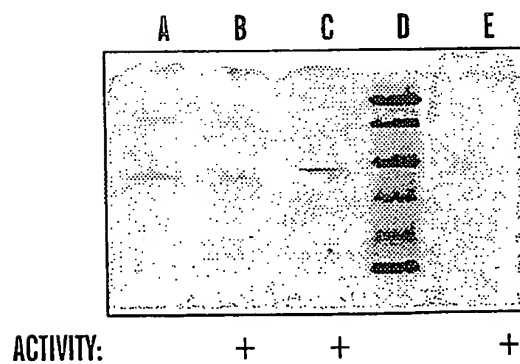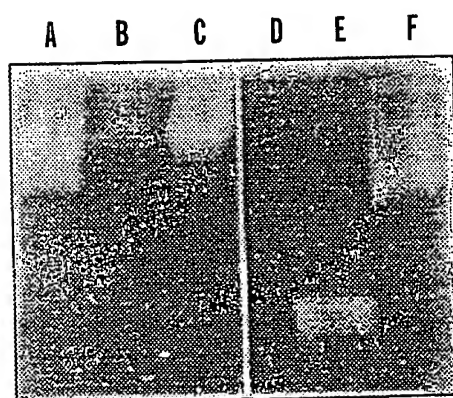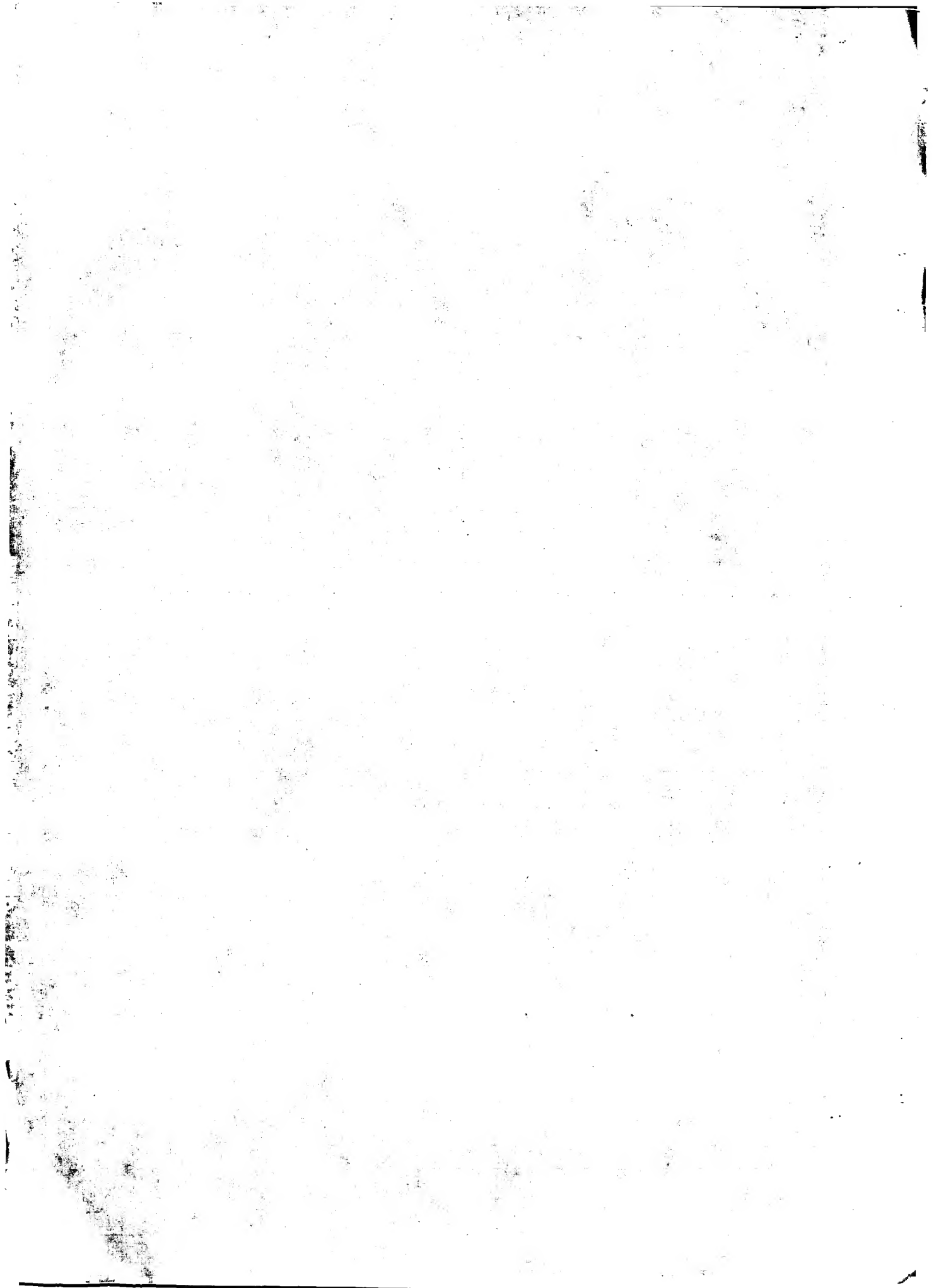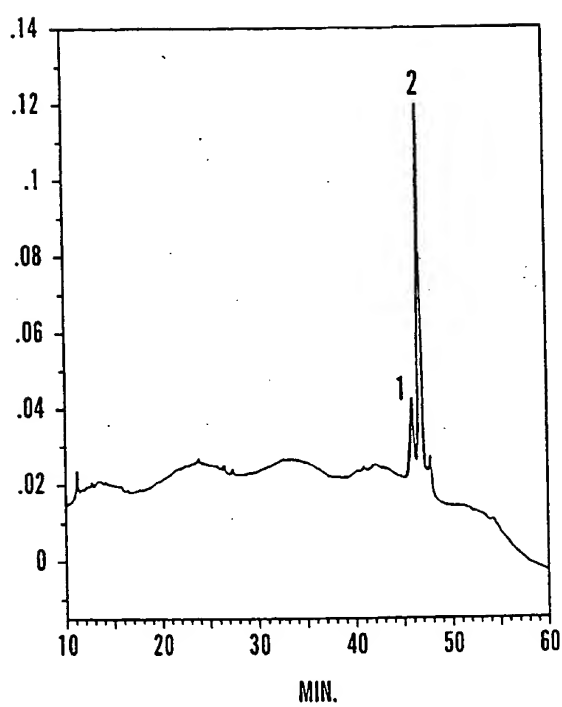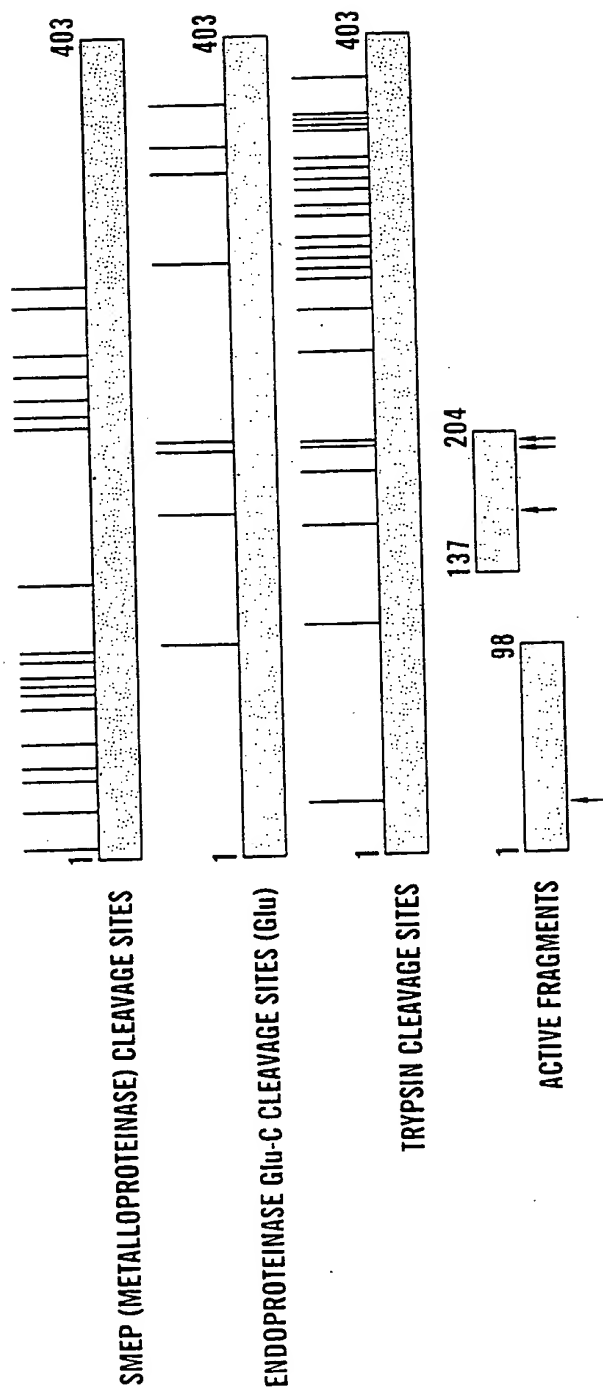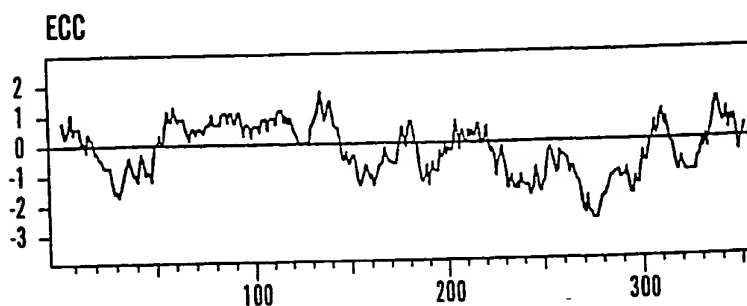
FIG. 7

Gene   Organism[a]   N-terminal amino acid sequence

HrpN   Ech   MQITIKAHIGGDLGVSGLGAQGLKGLNSAA
HrpN   Ecc   M.LNSLGGGAS.LQITIK.AGGNGGLFPSQ
HrpN   Eam   MSLNTSGLGASTMQISTGGAGGNNGLLGTS
WtsN   Est   MSMNTSPLGTSALQVTL...GGNNGLMGTD
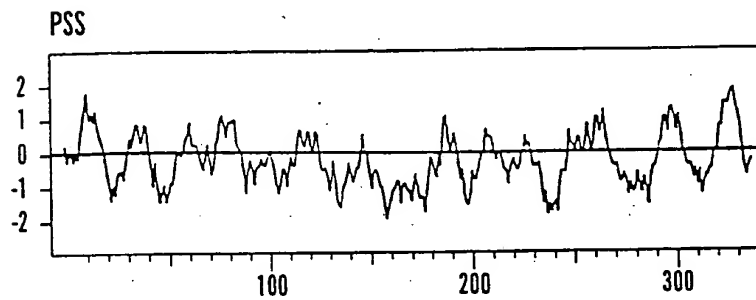orf1B   Eam   MSILTLNNNTSS.SPGLFQSGGDNGLGGHA

[a]Ech= E. chrysanthemi; Ecc= E. carotovora subsp. carotovora;
Eam= E.amylovora; Est= E. stewartii.

FIG. 8

EA

FIG. 9A

EST

FIG. 9B

ECH

FIG. 9C

ECC

FIG. 9D

11/13

RS

FIG. 9E

PSS

FIG. 9F

FIG. 10

N1;   5'-GGGAATTCATATGAGTCTGAATACAAGTGGG-3'
N76;   5'-GGGAATTCATATGGGCGGTGGCTTAGGCGGT-3'
N99;   5'-GGCATATGTCGAACGCGCTGAACGATATG-3'
N105;   5'-GGGAATTCATATGTTAGGCGGTTCGCTGAAC-3'
N110;   5'-GGCATATGCTGAACACGCTGGGCTCGAAA-3'
N137;   5'-GGCATATGTCAACGTCCCAAAACGACGAT-3'
N150;   5'-GGCATATGTCCACCTCAGACTCCAGCG-3'
N169;   5'-GGGAATTCATATGCAAAGCCTGTTTGGTGATGGG-3'
N210;   5'-GGGAATTCATATGGGTAATGGTCTGAGCAAG-3'
N267;   5'-GGGAATTCATATGAAAGCGGGCATTCAGGCG-3
N343;   5'-GGGAATTCATATGACACCAGCCAGTATGGAGCAG-3'
C75;   5'-GCAAGCTTAACAGCCCACCACCGCCCATCAT-3'
C104;   5'-GCAAGCTTAAATCGTTCAGCGCGTTCGACAG-3'
C168;   5'-GCAAGCTTAATATCTCGCTGAACATCTTCAGCAG-3'
C180;   5'-GCAAGCTTAAGGTGCCATCTTGCCCATCAC-3'
C204;   5'-GCAAGCTTAAATCAGTGACTCCTTTTTTATAGGC-3
C209;   5'-GCAAGCTTAACAGGCCCGACAGCGCATCAGT-3'
C266;   5'-GCAAGCTTAAACCGATACCGGTACCCACGGC-3'
C342;   5'-GCAAGCTTAATCCGTCGTCATCTGGCTTGCTCAG-3'
C403;   5'-GCAAGCTTAAGCCGCGCCCAGCTTG-3'

FIG. 11

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C12N15/31   C07K14/27   A01H1/00   A01H3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   C12N   C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 39802 A (CORNELL RES FOUNDATION INC) 19 December 1996 see abstract see page 10, line 30 – page 11, line 17 see page 12, line 21 – page 14, line 33 see page 17, line 33 – page 19, line 51 see page 27, line 11 – line 25 see page 29, line 28 – line 36 see examples 1–9 see figures 1,2 --- -/--	1-3, 8-10, 15-28, 35-37

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 November 1998	28/12/1998

Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Panzica, G

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category°	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 612 848 A (SANDOZ AG ;SANDOZ LTD (CH); SANDOZ AG (DE)) 31 August 1994 see page 2, line 28 - line 35 see page 2, line 51 - line 54 see page 4, line 23 - line 30	1-3,8-10
A	see page 5, line 12 - line 18	4-7, 11-43

X	WO 94 26782 A (CORNELL RES FOUNDATION INC) 24 November 1994	1-3, 8-10, 26-28
A	see page 2, line 11 - line 25 see page 3, line 14 - line 28	4-7, 11-14

X	WO 94 01546 A (CORNELL RES FOUNDATION INC) 20 January 1994	1-3, 8-10, 15-28, 35-37
A	see abstract see page 3, line 9 - page 4, line 10 see examples 1-15	4-7, 11-14

A	WO 93 23532 A (RIBOZYME PHARM INC) 25 November 1993 see abstract see page 1, line 23 - page 2, line 8 see page 2, line 25 - page 3, line 5	26-28, 35-37

X	WEI Z.M. ET AL.: "Harpin, elicitor of the hypersensitive response produced by the plant pathogen Erwinia amylovora" SCIENCE, vol. 257, 3 July 1992, pages 85-88, XP002085274 LANCASTER, PA US cited in the application see the whole document	1-3,8

X	BAUER D.W. ET AL.: "Erwinia chrysanthemi Harpin(Ech): an elicitor of the hypersensitive response that contributes to soft-rot pathogenesis" MOLECULAR PLANT-MICROBE INTERACTIONS, vol. 8, no. 4, 1995, pages 484-491, XP002085275 cited in the application	1-3,8
A	see the whole document	4-7, 11-14

	-/--	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate. of the relevant passages	Relevant to claim No.
X	WEI Z.M., BEER S.V.: "hrpL activates Erwinia amylovora hrp gene transcription and is a member of the ECF subfamily of "sigma" factors" JOURNAL OF BACTERIOLOGY, vol. 177, no. 21, November 1995, pages 6201-6210, XP002085276 see abstract	1-3,8
A	see figure 7	4-7, 11-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9639802	A	19-12-1996	US	5650387 A	22-07-1997
			AU	5982196 A	30-12-1996
			CA	2223616 A	19-12-1996
			CN	1192647 A	09-09-1998
			EP	0871354 A	21-10-1998
			FI	974430 A	26-01-1998
			PL	323823 A	27-04-1998
			US	5776889 A	07-07-1998
EP 0612848	A	31-08-1994	AU	682674 B	16-10-1997
			AU	5631194 A	01-09-1994
			CA	2116202 A	25-08-1994
			CZ	9400404 A	12-07-1995
			HU	69614 A	28-09-1995
			JP	6319558 A	22-11-1994
			PL	302320 A	05-09-1994
			SK	20694 A	10-05-1995
			US	5550228 A	27-08-1996
			US	5552527 A	03-09-1996
			ZA	9401281 A	24-08-1995
WO 9426782	A	24-11-1994	US	5708139 A	13-01-1998
			EP	0702694 A	27-03-1996
			JP	8510127 T	29-10-1996
WO 9401546	A	20-01-1994	EP	0648266 A	19-04-1995
			JP	7509604 T	26-10-1995
WO 9323532	A	25-11-1993	AU	4234293 A	13-12-1993
			CA	2135643 A	25-11-1993
			EP	0641384 A	08-03-1995
			JP	7507931 T	07-09-1995
			MX	9302804 A	29-04-1994